Distributed estimation for nonlinear systems ⋆

Stefano Battilotti [a] Matteo Mekhail [b]

[a] *Dipartimento di Ingegneria Informatica, Automatica e Gestionale "Antonio Ruberti" (e-mail: battilotti@dis.uniroma1.it).*

[b] *Dipartimento di Ingegneria Informatica, Automatica e Gestionale "Antonio Ruberti" (e-mail: mekhail@dis.uniroma1.it).*

Abstract

In a classical distributed framework, we present a novel distributed observer for genuinely nonlinear continuous-time plants. A network of sensors monitors a multiple-outputs plant. Each sensor measures only a portion of the plant's outputs and the sensing capability is different from sensor to sensor. The assumption of strongly connected digraph on the underlying sensor network ensures robustness and direct communication paths between nodes. Moreover, incremental homogeneity assumptions on the plant embrace a very large class of nonlinear systems for which a distributed observer can be designed. The distributed observer consists of local observers associated with each sensor, asymptotically estimating the entire state of the plant only by using the local sensing capability and information exchanged through the communication network. Numerical simulations on a network of interconnected Van Der Pol oscillators confirms theoretical results. Robustness and switching topologies are also discussed and suitable modifications of the distributed observer are proposed.

Key words: directed graphs, omniscience, distributed estimation, nonlinear systems

1 Introduction

The distributed state estimation has received an increasing attention and up-to-date it has been addressed with different approaches. Such attention is motivated by the feeling that, in many practical engineering systems, not a sufficient number of measurements can be made at a single location to give an asymptotic estimate of the plant state and the use of centralized techniques could be expensive or unfeasible. Loosely speaking, the framework of the distributed estimation problem considers a multiple-outputs plant monitored by a network of sensors. Each sensor measures only a portion of the output vector and the sensing capability is different from sensor to sensor. The main problem is to design a distributed observer associated with the sensors network, which estimate asymptotically the entire state of the plant only by the use of the local measurements and estimations exchanged over the communication network. The main challenge is represented by the local lack of observability of the main plant, i.e. from the sensor's perspective the plant is not observable. This implies that classical observer design techniques cannot be applied directly for each sensor. In the last decade different distributed techniques have been introduced, an overview of which can be found in [13]. There is a growing number of results in the observer-based design for both discrete and continuous time plants. The state of the art relies mostly on the joint (or collective) observability of the main plant and the definition of asymptotic omniscience in [16]. Such definition formalizes the general concept of agreement (or consensus) previously used in e.g. [14] to refer to an asymptotic decay of the estimation error in a distributed framework.

Distributed observers achieving asymptotic omniscience for discrete LTI systems are studied in [10][13]. A standing result is represented by [13] where authors have determined necessary and sufficient conditions for the existence of a reduced-order distributed observer satisfying certain scalability conditions. It should be noted that in this approach a sort of joint observability is addressed as observability by a set of source component representative nodes. In [10] a distributed observer is provided, extending the idea of the Kalman observable canonical decomposition to a setting with multiple sensor.
Another interesting approach [1] concerns the decentralized observability of a network, where methods and conditions on underlying graph is investigated to ensure joint observability even relatively to a subset of sensors.

For continuous-time results, the joint observability covers a crucial aspect in the design. Inspired by [16],

⋆ This work was supported by MIUR. This paper was not presented at any IFAC meeting.

[19, 8, 3, 21] provide conditions for the existence of omniscience-achieving Luenberger observers for continuous LTI plants with different underlying graphs. The procedure in [3], [8] relies both on orthogonal transformation which led to a Kalman decomposition. Moreover, in [3] strongly connected digraphs are handled with graph mirroring [15] while observer gains are computed by solving LMIs. In [19] the distributed stabilization problem has been accounted for LTI plants with a number of network topologies. Our previous work [9] studies the problem of achieving the asymptotic omniscience for nonlinear feedback linearizable systems, introducing the notion of semi-global omniscience (i.e. compact set of state trajectories to be estimated) and a semi-global omniscience-achieving observer is designed.

Robustness issues are discussed for linear models in [18] with $\mathcal{L}_2$ disturbances by using a $\mathcal{H}_\infty$ approach and resiliency to faults and attacks in [4] and [11] (with small delays). Time-varying topologies are considered for linear models in [18] (Markov-type switchings) and in [22] (deterministic-type switchings) with no disturbances: it is proved that if the dwell time of the switching law (i.e. the time interval between each switching and the next one) is long enough it is alway possible to obtain omniscience (in quadratic mean sense, wherever it applies).

In this paper we present a new class of omniscience-achieving observers for nonlinear systems under very general assumptions. We allow the presence of incremental homogeneous nonlinearities in the system's dynamics and its outputs. The systems we consider in this paper include a large class of nonlinear systems not limited to globally Lipschitz, lower/upper triangular or homogeneous systems. Moreover, the plant's dynamics is assumed to be bounded, which is reasonable even in a non-distributed context. As well as [13, 3, 9], we assume joint (or collective) observability of the plant. We borrow the definition of asymptotic omniscience from [13] and the distributed framework from [3], in particular we assume strongly connected and directed graphs on the underlying sensor network. Also, the present work generalizes in many directions our preliminary results in [9]. The novelty of our distributed observer is its nonlinear structure, which is introduced to cope with the nonlinearities of the plant.

Next we focus our attention on the robustness issue by introducing in the model $\mathcal{L}_\infty$ (i.e bounded time-varying) disturbances: we prove how to modify our distributed observer to take into account uncertainties/faults modeled by the disturbances. In this case, we formulate a robust version of the omniscience problem and under some conditions on the nonlinearities/uncertainties of the model, we prove that it is possible to achieve robust omniscience. Robust omniscience allows for a maximum tolerated estimation error (given by the designer) at each node (see [4], [11] for comparisons). $\mathcal{L}_2$ disturbances may be also taken into account by modifying the distributed observer according to a classical $\mathcal{H}_\infty$ approach. Our further contribution is to consider also time-varying topologies (changes in neighbors, in the number of assigned data packets per node and so on) and give conditions on the dwell time of the switching law under which a suitable modification of our distributed observer achieves robust omniscience (see [22], [18] for comparisons).

2 Notation

(**N1**) $\mathbb{R}^n$ (resp. $\mathbb{R}^{n\times n}$) is the set of n-dimensional real column vectors (resp. $n\times p$ matrices). $\mathbb{R}_{\geqslant}$ (resp. $\mathbb{R}^n_{\geqslant}$, $\mathbb{R}^{n\times p}_{\geqslant}$) denotes the set of real non-negative numbers (resp. vectors in $\mathbb{R}^n$, matrices in $\mathbb{R}^{n\times p}$, with real non-negative entries). $\mathbb{R}_{>}$ (resp. $\mathbb{R}^n_{>}$) denotes the set of real positive numbers (resp. vectors in $\mathbb{R}^n$ with real positive entries). $(\mathbb{R}^n)^*$ is the dual space of $\mathbb{R}^n$ (space of row vectors).
(**N2**) For any matrix $A\in\mathbb{R}^{p\times n}$ we denote by $A_{i,j}$ or $[A]_{i,j}$ the (i,j)-th entry of A and for any vector $v\in\mathbb{R}^n$ we denote by v_i the i-th element of v. Also, we may write vectors $v\in\mathbb{R}^n$ as $(v_1,\dots,v_n)^T$, vectors $w\in(\mathbb{R}^n)^*$ as $(w_1,\dots,w_n)$ and matrices $A\in\mathbb{R}^{s\times n}$ as $A=[v_1,\dots,v_n]$ (i.e. by columns) or $A=[w_1^T,\dots,w_s^T]^T$ (i.e. by rows). $\mathrm{Tr}\{A\}$ denotes the trace of $A\in\mathbb{R}^{n\times n}$. Moreover,

$$\mathrm{diag}\{A^{(1)},\dots,A^{(m)}\}=\begin{pmatrix} A^{(1)} & \cdots & 0\\ \vdots & \cdots & \vdots\\ 0 & \cdots & A^{(m)}\end{pmatrix}$$

where $A^{(j)}$ is any matrix and the 0 blocks have suitable dimensions. We retain a similar notation for functions. Also, $|a|$ denotes the absolute value of $a\in\mathbb{R}$, $\|a\|$ denotes the euclidean norm of $a\in\mathbb{R}^n$ with $\|a\|_M:=\sqrt{a^TMa}$, M positive semi-definite matrix, $\|A\|$ denotes the norm of $A\in\mathbb{R}^{n\times n}$ induced from the euclidean norm $\|\cdot\|$.
(**N3**) $\mathcal{K}$ denotes the set of continuous functions $f:\mathbb{R}_{\geqslant}\to\mathbb{R}_{\geqslant}$ strictly increasing and such that $f(0)=0$, $\mathcal{K}_\infty$ denotes the set of functions $f\in\mathcal{K}$ such that $f(s)\to+\infty$ as $s\to+\infty$.
(**N4**) A *saturation function σ_h with saturation levels* $h\in\mathbb{R}^n_{>}$ is a function $\sigma_h(x):=(\sigma_{h_1}(x_1),\dots,\sigma_{h_n}(x_n))^T$, $x\in\mathbb{R}^n$, such that for each $i=1,\dots,n$ and $x_i\in\mathbb{R}$:

$$\sigma_{h_i}(x_i)=\begin{cases} x_i & |x_i|\leqslant h_i\\ \mathrm{sign}(x_i)h_i & \text{otherwise.}\end{cases}\tag{1}$$

It is easy to prove that $|\sigma_{h_i}(x_i')-\sigma_{h_i}(x_i'')|\leqslant 2|\sigma_{h_i}(x_i'-x_i'')|$ and $|\sigma_{h_i}(x_i')|\leqslant h_i$ for each $h_i>0$ and for all $x_i',x_i''\in\mathbb{R}$.
(**N5**) For $\epsilon\in\mathbb{R}_{>}$, the group of dilations $\mathcal{G}=(\epsilon^{\mathfrak{r}},\diamond)$ is the set of elements $\epsilon^{\mathfrak{r}}:=(\epsilon^{\mathfrak{r}_1},\dots,\epsilon^{\mathfrak{r}_n})^T\in\mathbb{R}^n$, $\mathfrak{r}\in\mathbb{R}^n$, with group operation $\epsilon^{\mathfrak{r}'}\diamond\epsilon^{\mathfrak{r}''}=\epsilon^{\mathfrak{r}'+\mathfrak{r}''}$ and identity element $\mathbf{1}:=(1,\dots,1)^T$ (or $\mathbf{1}_n$ if we want to stress the dimension of the vector).

Also, we define the $\epsilon^{\mathfrak{r}}-$*dilation of* $v \in \mathbb{R}^n$ as the left group action $\diamond$ on $\mathbb{R}^n$ defined as $\epsilon^{\mathfrak{r}} \diamond v \coloneqq (\epsilon^{\mathfrak{r}_1} v_1, \cdots, \epsilon^{\mathfrak{r}_n} v_n)^T$. Similarly, we define the $\epsilon^{\mathfrak{r}}-$*dilation of* $w \in (\mathbb{R}^n)^*$ as the right group action $\diamond$ on $(\mathbb{R}^n)^*$ defined as $w \diamond \epsilon^{\mathfrak{r}} \coloneqq (\epsilon^{\mathfrak{r}_1} w_1, \cdots, \epsilon^{\mathfrak{r}_n} w_n)$.

By extension, we can define the *left* $\epsilon^{\mathfrak{r}}-$*dilation of* $A = [v_1^T, \ldots, v_n^T]^T \in \mathbb{R}^{n \times s}$ as the left group action $\diamond$ on $\mathbb{R}^{n \times s}$ defined as $\epsilon^{\mathfrak{r}} \diamond A \coloneqq [\epsilon^{\mathfrak{r}_1} v_1^T, \cdots, \epsilon^{\mathfrak{r}_n} v_n^T]^T$ and the *right* $\epsilon^{\mathfrak{r}}-$*dilation of* $A = [v_1, \ldots, v_n] \in \mathbb{R}^{s \times n}$ as the right group action $\diamond$ on $\mathbb{R}^{s \times n}$ defined as $A \diamond \epsilon^{\mathfrak{r}} \coloneqq [\epsilon^{\mathfrak{r}_1} v_1, \cdots, \epsilon^{\mathfrak{r}_n} v_n]$. Properties are given in the appendix.

(**N6**) for any vectors $x, y \in \mathbb{R}^n$ we write $x \preceq y$ if and only if $x_i \leqslant y_i$ for all $i = 1, \ldots, n$. We retain the same notation for matrices $A, B \in \mathbb{R}^{n \times n}$: $A \preceq B$ if and only if $A_{ij} \leqslant B_{ij}$ for all $i, j = 1, \ldots, n$.

3 The structure of the network

In this paper we consider *weighted* graphs denoted by $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \mathcal{A})$, $\mathcal{N}$ is a finite nonempty set of nodes, $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ is an edge of ordered pairs of nodes, and $\mathcal{A} \in \mathbb{R}^{N \times N}$ denotes the *adjacency matrix*, with N the cardinality of $\mathcal{N}$ (we will identify $\mathcal{N}$ with the set $\{1, \ldots, N\}$). The (i,j)-th entry $\mathcal{A}_{i,j}$ is the weight associated with the edge (i,j). We have $\mathcal{A}_{i,j} \neq 0$ if and only if $(i,j) \in \mathcal{E}$. Otherwise $\mathcal{A}_{i,j} = 0$. An edge $(i,j) \in \mathcal{E}$ means that the information flows from node i to node j. A graph is said to be *undirected* if it has the property that $(i,j) \in \mathcal{E} \Rightarrow (j,i) \in \mathcal{E}$ for all $i, j \in \mathcal{N}$; otherwise, we will say that the graph is *directed*. We will assume that the graph is *simple*, i.e. $\mathcal{A}_{i,i} = 0$ for all $i \in \mathcal{N}$. For an edge (i,j) node j is a *neighbor* of node i. A directed path from node i_1 to node i_l is a sequence of edges $(i_k, i_{k+1}), k = 1, 2, \ldots, l-1$. A directed graph $\mathcal{G}$ is *strongly connected* if between any pair of distinct nodes i and j in $\mathcal{G}$, there exists a directed path from i to j, $i, j \in \mathcal{N}$. In this paper we assume that the graph $\mathcal{G}$ is strongly connected.

The *Laplacian* $\mathcal{L} \in \mathbb{R}^{N \times N}$ is defined as $\mathcal{L} \coloneqq \mathcal{M} - \mathcal{A}$ where the i-th diagonal entry of the diagonal matrix $\mathcal{M}$ is given by $m_i = \sum_{j=1}^N \mathcal{A}_{i,j}$. By construction $\mathcal{L}$ has a zero eigenvalue with an associated eigenvector $\mathbf{1}_N$ (i.e. such that $\mathcal{L}\mathbf{1}_N = 0$) and if the graph is strongly connected all the other eigenvalues lie in the open right-half complex plane. For strongly connected graphs $\mathcal{G}$ it is possible to find a diagonal positive definite matrix D with positive elements such that $\hat{\mathcal{L}} = D\mathcal{L} + \mathcal{L}^T D$ is positive semi-definite (see lemma B.1). The matrix $D\mathcal{L}$ is the Laplacian of the balanced digraph obtained by adjusting the weights in the original graph. The matrix $\hat{\mathcal{L}}$ is the Laplacian of the undirected graph obtained by taking the union of the edges and their reversed edges in this balanced digraph. This undirected graph is called the *mirror* of this balanced graph.

4 The class of systems, problem statement and main assumptions

We consider continuous-time nonlinear systems:

$$\dot{x}_t = Ax_t + \phi(x_t),\ y_t = Cx_t + \psi(x_t), t \geqslant 0, \tag{2}$$

with state $x_t \in \mathbb{R}^n$, measurements $y_t \in \mathbb{R}^p$ and (C, A) observable. We assume that ϕ and ψ are locally Lipschitz and A, C of the form

$$\begin{aligned} &A = \operatorname{diag}\{A^{(1)}, \ldots, A^{(N)}\}, C = \operatorname{diag}\{C^{(1)}, \ldots, C^{(N)}\}, \\ &A^{(i)} = \operatorname{diag}\{A^{(i,1)}, \ldots, A^{(i,p^{(i)})}\}, \\ &C^{(i)} = \operatorname{diag}\{C^{(i,1)}, \ldots, C^{(i,p^{(i)})}\}, \end{aligned} \tag{3}$$

and $A^{(i,j)} \in \mathbb{R}^{n^{(i,j)} \times n^{(i,j)}}$ and $C^{(i,j)} \in \mathbb{R}^{1 \times n^{(i,j)}}$ having the form

$$A^{(i,j)} = \begin{pmatrix} 0 & 1 & 0 \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 \cdots & 1 \\ 0 & 0 & 0 \cdots & 0 \end{pmatrix},\ C^{(i,j)} = \begin{pmatrix} 1 & 0 \cdots & 0 \end{pmatrix},$$

with $\sum_{i=1}^N p^{(i)} = p$ and $\sum_{i=1}^N \sum_{j=1}^{p^{(i)}} n^{(i,j)} = n$. We define also the following set of indexes, used in our assumptions:

$$\begin{aligned} S = \{p^{(s,r)} \in \mathbb{N} : p^{(s,r)} = &\sum_{j=1}^{s} n^{(r+1,j)} + \sum_{i=1}^{r} \sum_{j=1}^{p^{(i)}} n^{(i,j)}, \\ &r = 0, \ldots, N-1, s = 1, \ldots, p^{(r+1)}\}. \end{aligned} \tag{4}$$

This is the set of indexes which correspond to the last row of each $A^{(i,j)}$ in A. Accordingly, y and $\psi(x)$ will be partitioned as

$$y = [{y^{(1)}}^T, \ldots, {y^{(N)}}^T]^T, \psi(x) = [{\psi^{(1)}}^T(x), \ldots, {\psi^{(N)}}^T(x)]^T.$$

and C by block rows as

$$C = [{H^{(1)}}^T, \ldots, {H^{(N)}}^T]^T \tag{5}$$

(i.e. $H^{(i)}$ represents the i-th block row of C: in our analysis we use this block decomposition of C rather than the one in (3) which is introduced only for defining precisely the block structure of C). Here, the portion $y^{(i)} \in \mathbb{R}^{p^{(i)}}$ is assumed to be the only information that can be processed by the node i. By assumption the pair (C, A) is observable (*joint or collective observability*) while each pair $(C^{(i)}, A)$ separately is not observable (*local unobservability*). For clarity, throughout the paper subscripts $_i$ will be always associated to components of vectors or functions and superscripts $^{(i)}$ will be associated to node variables.

4.1 Comments on the class of systems

For a linear system $\dot{z}_t = Jz_t,\ y_t = Wz_t$, with full row rank matrix W and (W, J) observable, there exists a change of coordinates $x = Tz$ for which in the new coordinates we have

$$\dot{x}_t = (A + BF)x_t,\ y_t = Cx_t, \tag{6}$$

where (C, A) is as in (3), B is

$$\begin{aligned} &B = \text{diag}\{B^{(1)}, \dots, B^{(N)}\}, \\ &B^{(i)} = \text{diag}\{B^{(i,1)}, \dots, B^{(i,p^{(i)})}\} \\ &B^{(i,j)} = \begin{pmatrix} 0 & 0 & \cdots & 1 \end{pmatrix}^T \in \mathbb{R}^{n^{(i,j)} \times 1}, \end{aligned} \tag{7}$$

and $F \in \mathbb{R}^{p \times n}$ ((6) is an observable canonical form for a linear multi-output system). The system (6) has the form (2) with $\phi(x) = BFx$ and $\psi(x) = 0$. Therefore, there is no loss of generality in considering the system in the form (2) (dumping the system nonlinearities in ψ and ϕ). The model (2) is enough general to include many nonlinear systems with observable linearization (in particular, globally Lipschitz systems when $\phi(x)$ is globally Lipschitz, triangular systems when $\phi(x)$ has a triangular structure, homogeneous systems when $\phi(x)$ is homogenous, Lur'e systems when $\phi(x) = \widetilde{\phi}(Cx)$ and so on). The reason for considering preferred coordinates for (2) is that in this coordinates it is possible to design nonlinear observers using suitable assumptions on the nonlinearities ϕ and ψ. This is a crucial step in the design of our distributed observer for (2) and for this reason we consider preferred coordinates (2). More generally, detectability of (C, A) in (2) can be taken into account by decomposing (2) into observable (C_O, A_O) (of the form (3)) and unobservable (C_U, A_U) subsystems and modify accordingly the distributed observer.

Notice that in the form (2),(3) orthogonal matrices $H^{(i)}$, $i = 1, \dots, N$, (the block rows of C in (5)) are required. This corresponds to assuming no overlapping measurements among nodes/sensors. In remark 5.1 we will point out how to take into account a possible overlapping of the measurements (i.e. $y^{(j)}$ and $y^{(i)}$ for some $j \neq i$ have at least one component in common) in the design of the distributed observer.

A more challenging (and more realistic) model than (2) is

$$\dot{x}_t = Ax_t + \phi(x_t, \delta_t),\ y_t = Cx_t + \psi(x_t, \delta_t), \tag{8}$$

$\delta_t \in \mathbb{R}^s$ a (time-varying) bounded disturbance. The disturbance δ_t models uncertainties of the system as well as faults. Another way of modeling δ_t is as a (time-varying) $\mathcal{L}_2$ (i.e. square integrable) disturbance (see for instance [18]). We will consider (8) in section 6 for discussing robustness issues.

4.2 Distributed estimation problem formulation and structure of the distributed observer

We will design a distributed nonlinear observer for the system (2) with the given communication network $\mathcal{G}$. Let $\mathfrak{g} \in \mathbb{R}^n$, $\mathfrak{r} \in \mathbb{R}^n_>, \gamma, c, \epsilon \in \mathbb{R}_>$, diagonal positive definite $D \in \mathbb{R}^{N \times N}$ and $\Gamma \in \mathbb{R}^{n \times n}$ be design parameters. The distributed observer will consists of N local observers and the local observer at the sensing node i has the following structure

$$\begin{aligned} \dot{\hat{x}}_t^{(i)} &= A\hat{x}_t^{(i)} + \phi\left(\sigma_{c\epsilon^{\mathfrak{r}}}(\hat{x}_t^{(i)})\right) \\ &\quad + \frac{\text{Tr}(D)}{D_{i,i}} L^{(i)} Z_t^{(i)} + \gamma P^{-1} K_t^{(i)},\ i \in \mathcal{N},\ \hat{x}_0^{(i)} = 0, \end{aligned} \tag{9}$$

with innovation process

$$Z_t^{(i)} := y_t^{(i)} - H^{(i)}\hat{x}_t^{(i)} - \psi^{(i)}\left(\sigma_{c\epsilon^{\mathfrak{r}}}(\hat{x}_t^{(i)})\right),$$

consensus term

$$K_t^{(i)} := \sum_{j=1}^{N} \mathcal{A}_{i,j}(\hat{x}_t^{(j)} - \hat{x}_t^{(i)}) \tag{10}$$

and matrices

$$\begin{aligned} &L^{(i)} := P^{-1} {H^{(i)}}^T R^{(i)}, P := (I - GA^T)^T \diamond \epsilon^{-2\mathfrak{r}} \diamond (I - GA^T), \\ &G := \epsilon^{\mathfrak{g}} \diamond \Gamma \diamond \epsilon^{\mathfrak{g}},\ R^{(i)} := H^{(i)}(\epsilon^{-\mathfrak{r}} \diamond G \diamond \epsilon^{-\mathfrak{r}}){H^{(i)}}^T, \end{aligned} \tag{11}$$

where $\hat{x}^{(i)}$ is the state of the local observer at the sensing node i and $\mathcal{A}_{i,j}$ is the (i, j) entry of the adjacency matrix $\mathcal{A}$ of the given network $\mathcal{G}$. Each local observer uses the local innovation $Z^{(i)}$ (obtained from the partial information y_i available at the node) and exchanges its state estimate with the state estimates of its neighbors. The local observers are initialized at $\hat{x}_0^{(i)} := 0$ but this has been done only for simplicity (this restriction can be easily relaxed: see remark 5.2).

Remark 4.1 *The distributed observer (9) is such that at each node the following information is processed: a local information (the local information on the network, the data packet $y^{(i)}$ processed at the sensing node and the neighbor's estimates $\hat{x}^{(j)}$, $(i, j) \in \mathcal{E}$) and a global information (the matrix D which, as it will be seen, defines the Laplacian $\hat{\mathcal{L}} = D\mathcal{L} + \mathcal{L}^T D$ of the mirror graph and the parameters γ, c, ϵ and Γ which rely on a perfect knowledge of the system nonlinearities $\phi(x)$ and $\psi(x)$ with a consequent lack of robustness). However, by definition (see lemma B.1) $N = \mathbf{1}_N^T D \mathbf{1}_N = \text{Tr}(D)$ and $\mathbf{1}_N^T D\mathcal{L} = 0$ which, by the structure of the Laplacian $\mathcal{L}$, implies that each $D_{i,i}$, $i = 1, \dots, N$, depends on the number N of the nodes and on the structure of the network around the node i (i.e. how it is connected to its neighbors). Therefore, $D_{i,i}$*

in (9) is rather a local information while $\mathrm{Tr}(D)$ *is equal to the number* N *of nodes in the network. With regard to the parameters* γ, c, ϵ *and* Γ*, in section 6 by considering the uncertain model (8), we will modify the distributed observer (9) in such a way that the parameters* γ, c, ϵ *and* Γ *will depend on the system nonlinearities* $\phi(x,0)$ *and* $\psi(x,0)$ *which contain much less information than* $\phi(x,\delta)$ *and* $\psi(x,\delta)$ *and guarantee the robustness of the distributed observer against disturbances/uncertainties* δ *(distributed observer (43)).*

The task of a distributed observer for (2) is to cooperatively estimate the state of the system (2). For a formulation of our problem we will refer to the notion of omniscience introduced in [13] with a slight generalization (the set $\mathcal{C}$ below is all the state space in [13]).

Definition 4.1 *(Asymptotic Omniscience (relatively to* $\mathcal{C} \subset \mathbb{R}^n$*)): A distributed observer* $\{\hat{x}^{(i)}\}_{i\in\mathcal{N}}$ *achieves asymptotic omniscience for (2) if for all the state trajectories* x_t *of (2) in* $\mathcal{C} \subset \mathbb{R}^n$ *we have* $\lim_{t\to+\infty} \|\hat{x}_t^{(i)} - x_t\| = 0$ *for all* $i \in \mathcal{N}$*, i.e. the state estimate* $\hat{x}_t^{(i)}$ *at each node* i *asymptotically converges to* x_t*.*

Given $\mathcal{C} \subset \mathbb{R}^n$, asymptotic omniscience (relatively to $\mathcal{C}$) is guaranteed for all the trajectories x_t in $\mathcal{C}$ (or for all initial conditions x_0 in a smaller set $\mathcal{C}_S \subseteq \mathcal{C}$): in comparison with [13], this is a kind of *semiglobal* (relatively to $\mathcal{C}$) omniscience property. Global omniscience ($\mathcal{C} = \mathbb{R}^n$) for nonlinear systems will be not considered in this paper, since it requires much stricter conditions on the nonlinearities.

4.3 Assumptions and main result

Our assumptions on the system (2) are the following (see a review of various notions of incremental homogeneity in appendix A).

(**H0**) (*state boundedness*): *the state trajectories* x_t *of (2) are contained in some known compact set* $\mathcal{C} \subset \mathbb{R}^n$*,*

(**H1**) (*incremental homogeneity in the upper bound*): *for some degrees* $\mathfrak{g}, \hat{\mathfrak{g}} \in \mathbb{R}^n$ *and weights* $\mathfrak{r} \in \mathbb{R}^n_>$ *such that for each* $j \notin S$

$$2\mathfrak{g}_{j+1} - \mathfrak{g}_j \leqslant \hat{\mathfrak{g}}_j \leqslant \mathfrak{g}_j, \tag{12}$$

where $\hat{\mathfrak{g}} \in \mathbb{R}^n$ *is defined component-wise as follows*

$$\hat{\mathfrak{g}}_i := \mathfrak{g}_i,\ i \in S;\ \hat{\mathfrak{g}}_i := \mathfrak{r}_{i+1} - \mathfrak{r}_i - \mathfrak{g}_{i+1},\ i \notin S, \tag{13}$$

(i) ϕ *is incrementally homogeneous in the upper bound with quadruples* $\{\mathfrak{r}, \mathfrak{r} + \hat{\mathfrak{g}}, \mathfrak{g}, \Phi(x', x'')\}$*, with* $\Phi(0,0) = BF$*,* B *as in (7) and for some* $F \in \mathbb{R}^{p\times n}$*,*

(i) $C^T\psi$ *is incrementally homogeneous in the upper bound with quadruples* $\{\mathfrak{r}, \mathfrak{r} - \mathfrak{g}, \mathfrak{g}, C^T\Psi(x', x'')\}$*, with* $\Psi(0,0) = \alpha C$ *and* $\alpha \in [0,1)$*,*

(**H2**) (*uniform incremental homogeneity in the* 0*-limit*): ϕ *is uniformly incrementally homogeneous in the* 0*-limit with triple* $\{\mathbf{1}_n, \mathbf{1}_n, \frac{\partial\phi}{\partial x'}(x')x''\}$ *and* $C^T\psi$ *is uniformly incrementally homogeneous in the* 0*-limit with triple* $\{\mathbf{1}_n, \mathbf{1}_n, C^T\frac{\partial\psi}{\partial x'}(x')x''\}$*.*

The main result of this paper concerns the achievement of (semiglobal) omniscience and it is stated as follows.

Theorem 4.1 *Consider a network* $\mathcal{N}$ *described by a directed strongly connected graph* $\mathcal{G}$ *and a system (2) with* (C, A) *observable (collective observability). Given a compact set* $\mathcal{C} \subset \mathbb{R}^n$ *and under assumptions (***H0***)-(***H2***) there exist* $\epsilon^* \geqslant 1$*, diagonal positive definite* Γ *and* D *and* $\gamma, c \in \mathbb{R}_>$ *such that for all* $\epsilon \geqslant \epsilon^*$ *the distributed observer* $\{\hat{x}^{(i)}\}_{i\in\mathcal{N}}$ *over the given network* $\mathcal{N}$ *described by (9) achieves asymptotic omniscience for (2) relatively to* $\mathcal{C}$*.*

4.4 Comments on the assumptions on the system

Assumptions (**H0**)-(**H1**) are enough general for coping with genuinely nonlinear systems and were adopted (with additional restrictive conditions) in [2] for designing observers for a single-output nonlinear system with observable linearization.

Assumption (**H0**) requires that the state trajectories of (2) live for all times inside a known compact set $\mathcal{C}$, which means that we are restricting the set of initial conditions x_0 to a known compact set $\mathcal{C}_S \subseteq \mathcal{C}$.

Assumption (**H1**) is met for a large class of nonlinear systems (2), in particular any system (2) with either one of: linear, globally Lipschitz, lower triangular, upper triangular or homogeneous $\phi(x)$. Assumption (**H1**) amounts to solve a set of algebraic inequalities in the unknowns $\mathfrak{r} \in \mathbb{R}^n_>$ and $\mathfrak{g} \in \mathbb{R}^n$ (see examples A.1 and A.2 in appendix A). In particular, notice the conditions $\Phi(0,0) = BF$ (the linearization of $\phi(x)$ around 0 has the form BFx: compare with the analogous term in (6)) and $\Psi(0,0) = \alpha C$ with $\alpha \in [0,1)$ (the linearization of $\psi(x)$ around 0 has the form αCx with $\alpha \in [0,1)$, which is a sector-like condition). For example

$$\phi(x) = \begin{pmatrix} 0 \\ -\mu x_1 + \mu x_2(1 - x_1^2) \\ 0 \\ ax_1 + bx_2 - \mu x_3 + \mu x_4(1 - x_3^2) \end{pmatrix}$$

(see section 8 for a more general example), with $\mu, a, b > 0$, satisfies (**H1**) with $\mathfrak{r}_1 = \mathfrak{r}_3 = 1$, $\mathfrak{r}_2 = \mathfrak{r}_4 = 3$ and

$\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}_3 = \mathfrak{g}_4 = 1$ (in this example $S = \{2, 4\}$), $B = \text{diag}\{(0,1)^T, (0,1)^T\}$ and

$$F = \begin{pmatrix} \mu & \mu & 0 & 0 \\ a & b & \mu & \mu \end{pmatrix}.$$

Assumption (**H2**) is always satisfied for polynomial functions ϕ and ψ and requires that small increments of ϕ and ψ from some given point $x \in \mathcal{C}$ are uniformly (over $\mathcal{C}$) approximated by their linear approximation around x.

A fact extensively used in the proofs of our results is that, for $\epsilon \geqslant 1$, we can (lower and upper) bound the action of a dilation $\epsilon^{\mathfrak{g}}$ on $\mathbb{R}^n$ as follows:

$$\epsilon^{\min_j\{\mathfrak{g}_i\}}\mathbf{1}_n \preceq \epsilon^{\mathfrak{g}} \preceq \epsilon^{\max_j\{\mathfrak{g}_i\}}\mathbf{1}_n. \tag{14}$$

5 Proof of the main theorem

The rationale behind the proof of theorem 4.1 is the following. Consider the estimation error e resulting from (2) and the N local observers (9) and change error coordinates $\hat{e}$ in such a way to diagonalize the Laplacian $\hat{\mathcal{L}}$ of the mirror graph. Associated to the Laplacian we have a potential $\mathcal{P}(\hat{e}) = \hat{e}^T\hat{\mathcal{L}}e$. With a Lyapunov function $\hat{V}(\hat{e})$ we obtain that the time derivative $\dot{\hat{V}}(\hat{e})$ along the estimation error trajectories is negative definite if the potential $\mathcal{P}(\hat{e})$ is zero. In particular, it is equal to the time derivative of a certain Lyapunov function along the error trajectories between the state trajectories of (2) and the estimate provided by a "centralized" observer (i.e. which processes all the data packets of the network). Using this and the fact that $\dot{\hat{V}}(\hat{e})$ is quadratic in $\hat{e}$ for small $\hat{e}$ (by (H2)), γ in (9) is designed so that $\dot{\hat{V}}(\hat{e})$ is negative definite also for nonzero potential $\mathcal{P}(\hat{e})$. The parameters c, ϵ and Γ are designed on the "centralized" observer (using (H0), (H1)). Asymptotic convergence of the estimation error $\hat{e}$ follows from Lyapunov theorems.

5.1 The estimation error dynamics and convergence properties

In this section we study some stability properties of the estimation error system resulting from (2) and the N local observers (9). Let the matrix D be selected as in lemma B.1, i.e. D is diagonal positive definite such that $D\mathcal{L} + \mathcal{L}^T D$ is positive semi-definite. Also, $\bar{c} > 0$ and $\Gamma \in \mathbb{R}^{n\times n}$ be selected as in lemma B.4 and pick $c \in (0, \bar{c})$. On account of (**H0**), we will assume that $\epsilon_1 > 1$ have been selected so that

$$\sigma_{c\epsilon^{\mathfrak{r}}}(x) = x \qquad \forall \epsilon \geqslant \epsilon_1,\ x \in \overline{\mathcal{B}(\mathcal{C})} \tag{15}$$

($\overline{\mathcal{B}(\mathcal{C})}$ is the closure of an open set $\mathcal{B} \supset \mathcal{C}$) in other words the saturation $\sigma_{c\epsilon^{\mathfrak{r}}}$ is not active on $\overline{\mathcal{B}(\mathcal{C})}$. From now on we consider $\epsilon \geqslant \epsilon_1$. We also write for simplicity $\sigma(\cdot)$ instead of $\sigma_{c\epsilon^{\mathfrak{r}}}(\cdot)$ and for notational convenience we use the following incremental operator Δ with any $f \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^n)$: $\Delta f(\chi_1, \chi_2) := f(\chi_1) - f(\chi_2)$.

If $e^{(i)} := x - \hat{x}^{(i)}$ and $e := ((e^{(1)})^T, \ldots, (e^{(N)})^T)^T$, on account of (15), we have with the Kronecker product formalism (and using the property $(A_1 \otimes A_2)(B_1 \otimes B_2) = (A_1B_1) \otimes (A_2B_2)$)

$$\begin{aligned}\dot{e}_t = &\big(I_N \otimes A - \text{Tr}\{D\}(D^{-1} \otimes I_n)LQ\big)\, e_t - \gamma(\mathcal{L} \otimes P^{-1})e \\ &+ F(x, e_t) - \text{Tr}\{D\}(D^{-1} \otimes I_n)LQU(x, e_t)\end{aligned} \tag{16}$$

with

$$L := \text{diag}\{L^{(1)}, \ldots, L^{(N)}\}, Q := \text{diag}\{H^{(1)}, \ldots, H^{(N)}\} \tag{17}$$

$$F(x,e) := \begin{pmatrix} \Delta(\phi \circ \sigma)(x, x - e^{(1)}) \\ \vdots \\ \Delta(\phi \circ \sigma)(x, x - e^{(N)}) \end{pmatrix}$$

$$U(x,e) := \begin{pmatrix} \Delta({H^{(1)}}^T\psi^{(1)} \circ \sigma)(x, x - e^{(1)}) \\ \vdots \\ \Delta({H^{(N)}}^T\psi^{(N)} \circ \sigma)(x, x - e^{(N)}) \end{pmatrix}.$$

Also, notice that on account of (H3), since by (15) and compactness of $\mathcal{C}$, for any compact set $\mathcal{S} \subset \mathbb{R}^n$ we have:

$$\lim_{\lambda\to 0} \max_{\substack{x\in\mathcal{C} \\ e^{(i)}\in\mathcal{S}}} \left\| \frac{\Delta(\phi \circ \sigma)(x, x - \lambda e^{(i)})}{\lambda} - \frac{\partial\phi}{\partial x}(x)e^{(i)} \right\| = 0$$

$$\lim_{\lambda\to 0} \max_{\substack{x\in\mathcal{C} \\ e^{(i)}\in\mathcal{S}}} \Big\| \frac{\Delta({H^{(i)}}^T\psi^{(i)} \circ \sigma)(x, x - \lambda e^{(i)})}{\lambda} - {H^{(i)}}^T\frac{\partial\psi^{(i)}}{\partial x}(x)e^{(i)} \Big\| = 0.$$

for each $i = 1, \ldots, N$. Therefore, if we define

$$F_0(x,e) := \text{diag}\left\{\frac{\partial\phi}{\partial x}(x), \cdots, \frac{\partial\phi}{\partial x}(x)\right\} e,$$

$$U_0(x,e) := \text{diag}\Big\{{H^{(1)}}^T\frac{\partial\psi^{(1)}}{\partial x}(x), \cdots, {H^{(N)}}^T\frac{\partial\psi^{(N)}}{\partial x}(x)\Big\}e,$$

we have the following useful property: for any N-tuple of compact sets $\mathcal{S}^{(1)}, \ldots, \mathcal{S}^{(N)} \subset \mathbb{R}^n$ with $\mathcal{S}^{\times N} := \mathcal{S}^{(1)} \times \cdots \times \mathcal{S}^{(N)}$

$$\begin{aligned}&\lim_{\lambda\to 0} \max_{\substack{x\in\mathcal{C} \\ e\in\mathcal{S}^{\times N}}} \left\| \frac{F(x, \lambda e)}{\lambda} - F_0(x,e) \right\| = 0 \\ &\lim_{\lambda\to 0} \max_{\substack{x\in\mathcal{C} \\ e\in\mathcal{S}^{\times N}}} \left\| \frac{U(x, \lambda e)}{\lambda} - U_0(x,e) \right\| = 0.\end{aligned} \tag{18}$$

For any invertible $R \in \mathbb{R}^{nN\times nN}$ and since R maps compact sets into compact sets, we also have

$$\lim_{\lambda\to 0}\max_{\substack{x\in\mathcal{C}\\ e\in\mathcal{S}\times N}}\left\|\frac{F(x,\lambda Re)}{\lambda}-F_0(x,Re)\right\|=0$$
$$\lim_{\lambda\to 0}\max_{\substack{x\in\mathcal{C}\\ e\in\mathcal{S}\times N}}\left\|\frac{U(x,\lambda Re)}{\lambda}-U_0(x,Re)\right\|=0. \tag{19}$$

Consider the following change of coordinates

$$\hat{e}=(T\otimes I_n)e \tag{20}$$

where $T\in\mathbb{R}^{N\times N}$ is an orthogonal matrix chosen as in lemma B.2. In these new coordinates (16) reads out as

$$\begin{aligned}\dot{\hat{e}}_t=&(T\otimes I_n)A_D(T^T\otimes I_n)\hat{e}_t-\gamma((T\mathcal{L}T^T)\otimes P^{-1})\hat{e}_t\\ &+(T\otimes I_n)F\left(x_t,(T^T\otimes I_n)\hat{e}_t\right)\\ &+\mathrm{Tr}\{D\}((TD^{-1})\otimes I_n)LQU\left(x_t,(T^T\otimes I_n)\hat{e}_t\right)\end{aligned} \tag{21}$$

where we set

$$A_D:=\left(I_N\otimes A-\mathrm{Tr}\{D\}(D^{-1}\otimes I_n)LQ\right).$$

Now, consider a candidate Lyapunov function for (21)

$$\hat{V}(\hat{e}):=\hat{e}^T(T\otimes I_n)(D\otimes P)(T^T\otimes I_n)\hat{e}. \tag{22}$$

We want to prove that the derivative of $\hat{V}$ along the trajectories of (21) is negative definite. Using (B.7) with orthogonality of T (lemma B.2) and the properties of the Kronecker product

$$\begin{aligned}&(T\otimes I_n)[(D\otimes P)(T^T\otimes I_n)((T\mathcal{L}T^T)\otimes P^{-1})\\ &+((T\mathcal{L}T^T)\otimes P^{-1})(T^T\otimes I_n)(D\otimes P)](T^T\otimes I_n)\\ &=\left[T(D\mathcal{L}+\mathcal{L}^TD)T^T\right]\otimes I_n=\Lambda\otimes I_n.\end{aligned} \tag{23}$$

From (21) and properties of the Kronecker product we obtain

$$\begin{aligned}\dot{\hat{V}}|_{(21)}=&-\gamma\hat{e}^T(\Lambda\otimes I_n)\hat{e}+2\Big[\hat{e}^T(TD\otimes P)A_D(T^T\otimes I_n)\hat{e}\\ &+\hat{e}^T((TD)\otimes P)F\left(x,(T^T\otimes I_n)\hat{e}\right)\\ &-\mathrm{Tr}\{D\}\hat{e}^T(T\otimes P)LQU\left(x,(T^T\otimes I_n)\hat{e}\right)\Big]\end{aligned} \tag{24}$$

Introduce the set

$$\mathcal{F}:=\{\hat{e}\in\mathbb{R}^{nN}:\hat{e}^T(\Lambda\otimes I_n)\hat{e}=0\}$$

and claim for a moment that

$$\begin{aligned}&\forall\hat{e}\in\mathcal{F}\backslash\{0\},\ \forall x\in\mathcal{C}\Rightarrow\\ &\mathcal{W}(x,\hat{e}):=2\Big[\hat{e}^T(TD\otimes P)A_D(T^T\otimes I_n)\hat{e}\\ &\quad+\hat{e}^T((TD)\otimes P)F\left(x,(T^T\otimes I_n)\hat{e}\right)\\ &\quad-\mathrm{Tr}\{D\}\hat{e}^T(T\otimes P)LQU\left(x,(T^T\otimes I_n)\hat{e}\right)\Big]<0\end{aligned} \tag{25}$$

and, in addition,

$$\begin{aligned}&\forall\hat{e}\in\mathcal{F}\backslash\{0\},\ \forall x\in\mathcal{C}\Rightarrow\\ &\mathcal{W}_0(x,\hat{e}):=2\Big[\hat{e}^T(TD\otimes P)A_D(T^T\otimes I_n)\hat{e}\\ &\quad+\hat{e}^T((TD)\otimes P)F_0\left(x,(T^T\otimes I_n)\hat{e}\right)\\ &\quad-\mathrm{Tr}\{D\}\hat{e}^T(T\otimes P)LQU_0\left(x,(T^T\otimes I_n)\hat{e}\right)\Big]<0.\end{aligned} \tag{26}$$

Notice that by construction

$$P\sum_{j=1}^{N}L^{(i)}H^{(i)}=C^TC(\epsilon^{-\mathfrak{r}}\diamond G\diamond\epsilon^{-\mathfrak{r}})C^TC \tag{27}$$

and

$$H^{(i)}{H^{(j)}}^T=0,\forall j\neq i. \tag{28}$$

Moreover, since $\Lambda=\mathrm{diag}\{0,\lambda_2,\dots,\lambda_N\}$ with $0<\lambda_2\leqslant\cdots\leqslant\lambda_N$ (lemma B.2), clearly

$$\hat{e}\in\mathcal{F}\Leftrightarrow\hat{e}_j=0,\ j=2,\dots,N.$$

On account of the above and since the first row of T is $\frac{1}{\sqrt{N}}\mathbf{1}_N^T$ (lemma B.2), it is straightforward to see that

$$\begin{aligned}&\forall\hat{e}\in\mathcal{F},\ \forall x\in\mathcal{C}\Rightarrow\mathcal{W}(x,\hat{e})=\\ &\mathcal{R}^{(g)}(x,\frac{\hat{e}_1}{\sqrt{N}}):=2(\frac{\hat{e}_1}{\sqrt{N}})^T\mathrm{Tr}\{D\}P\Big\{(A-KC)\frac{\hat{e}_1}{\sqrt{N}}\\ &\qquad+\Delta(\phi\circ\sigma)\left(x,x-\frac{\hat{e}_1}{\sqrt{N}}\right)\\ &\qquad-KC\Delta(C^T\psi\circ\sigma)\left(x,x-\frac{\hat{e}_1}{\sqrt{N}}\right)\Big\}\end{aligned} \tag{29}$$

and

$$\begin{aligned}&\forall\hat{e}\in\mathcal{F},\ \forall x\in\mathcal{C}\Rightarrow\\ &\mathcal{W}_0(x,\hat{e})=\mathcal{R}^{(L)}(x,\frac{\hat{e}_1}{\sqrt{N}}):=2(\frac{\hat{e}_1}{\sqrt{N}})^T\mathrm{Tr}\{D\}P\Big\{(A-KC)\\ &\qquad+\frac{\partial\phi}{\partial x}(x)-KCC^T\frac{\partial\psi}{\partial x}(x)\Big\}\frac{\hat{e}_1}{\sqrt{N}}\end{aligned} \tag{30}$$

where

$$\begin{aligned}&K=P^{-1}C^TR,\ P=(I-GA^T)^T\diamond\epsilon^{-2\mathfrak{r}}\diamond(I-GA^T),\\ &R=C(\epsilon^{-\mathfrak{r}}\diamond G\diamond\epsilon^{-\mathfrak{r}})C^T\end{aligned} \tag{31}$$

and G as in (11). Therefore, our claim (25) boils down to prove the existence of a *centralized* observer for (2) having the form

$$\dot{\hat{x}}_t^{(g)}=A\hat{x}_t^{(g)}+\phi(\sigma(\hat{x}_t^{(g)}))+KZ_t^{(g)} \tag{32}$$

with innovation process $Z_t^{(g)} := y_t - C\hat{x}_t^{(g)} - \psi\left(\sigma(\hat{x}_t^{(g)})\right)$ and matrices K, P and R as in (31) and G as in (11), using a Lyapunov function $V(\hat{e}^{(g)}) := \mathrm{Tr}\{D\}\hat{e}^{(g)T}P\hat{e}^{(g)}$ for the estimation error system with state $\hat{e}^{(g)} = x - \hat{x}^{(g)}$ (this will be accomplished in section 5.2). At this point, it comes into play the assumption of collective observability, i.e.(C, A) is an observable pair.

On the other hand, our second claim (26) boils down to prove for the system

$$\dot{x}_t^{(L)} = \left[A + \frac{\partial\phi}{\partial x_t}(x_t)\right] x_t^{(L)},\ y_t^{(L)} = Cx_t^{(L)} + \frac{\partial\psi}{\partial x_t}(x_t)x_t^{(L)} \tag{33}$$

(i.e. the system (2) linearized around its trajectory x_t) the existence of a *centralized* observer having the form

$$\dot{\hat{x}}_t^{(L)} = \left[A + \frac{\partial\phi}{\partial x_t}(x_t)\right] \hat{x}_t^{(L)} + KZ_t^{(L)} \tag{34}$$

with innovation process $Z_t^{(L)} := y_t^{(L)} - C\hat{x}_t^{(L)} - \frac{\partial\psi}{\partial x_t}(x_t)\hat{x}_t^{(L)}$ and matrices K, P and R as in (31) and G as in (11), using a Lyapunov function $V_L(\hat{e}^{(L)}) := \mathrm{Tr}\{D\}\hat{e}^{(L)T}P\hat{e}^{(L)}$ for the estimation error system with state $\hat{e}^{(L)} = x^{(L)} - \hat{x}^{(L)}$ (this will be accomplished in section 5.3). Collective observability, plays a key role also at this point.

Having in mind (25) and (26) and going back to (24), we are in a position to arrange the right-hand part of (24) to be negative definite with a suitable selection of $\gamma > 0$. Indeed, both (25) and (26) are crucial for the existence of $\gamma > 0$ for which (24) is negative (the first, for large values of the estimation error, the second for small values). This is direct consequence of lemma B.3 with $\Omega := \{\hat{e} \in \mathbb{R}^{nN} : \hat{V}(\hat{e}) \leqslant \omega\}$ for a fixed $\omega > 0$, $\mathcal{W}(x, \hat{e})$ in (25) and $\mathcal{W}_0(x, \hat{e})$ in (26). The functions $\mathcal{W}(x, \hat{e})$ and $\mathcal{W}_0(x, \hat{e})$ satisfy (B.8) by virtue of (19):

$$\begin{aligned}
&\lim_{\lambda\to 0} \max_{\substack{x\in\mathcal{C}\\ \hat{e}\in\mathcal{S}^{\times N}}} \left\| \frac{\mathcal{W}(x,\lambda\hat{e})}{\lambda^2} - \mathcal{W}_0(x,\hat{e}) \right\| \\
&\leqslant \max_{\substack{x\in\mathcal{C}\\ \hat{e}\in\mathcal{S}^{\times N}}} \left\| ((DT^T)\otimes P)\hat{e} \right\| \times \\
&\times \lim_{\lambda\to 0} \max_{\substack{x\in\mathcal{C}\\ \hat{e}\in\mathcal{S}^{\times N}}} \left\| \frac{F\left(x, \lambda(T^T\otimes I_n)\hat{e}\right)}{\lambda} - F_0\left(x, (T^T\otimes I_n)\hat{e}\right) \right\| \\
&+ \mathrm{Tr}\{D\} \max_{\substack{x\in\mathcal{C}\\ e\in\mathcal{S}^{\times N}}} \left\| Q^T L^T (T^T\otimes P)\hat{e} \right\| \times \\
&\times \lim_{\lambda\to 0} \max_{\substack{x\in\mathcal{C}\\ \hat{e}\in\mathcal{S}^{\times N}}} \left\| \frac{U\left(x, \lambda(T^T\otimes I_n)\hat{e}\right)}{\lambda} - U_0\left(x, (T^T\otimes I_n)\hat{e}\right) \right\| \\
&= 0.
\end{aligned}$$

From lemma B.3 we obtain the existence of $\gamma > 0$ such that

$$\forall \hat{e} \neq 0 : \hat{V}(\hat{e}) \leqslant \omega \Rightarrow \dot{\hat{V}}|_{(21)} < 0,\ \forall x \in \mathcal{C}. \tag{35}$$

This implies, with $V(e) := e^T(D\otimes P)e$ that:

$$\forall e \neq 0 : V(e) \leqslant \omega \Rightarrow \dot{V}|_{(16)} < 0,\ \forall x \in \mathcal{C}. \tag{36}$$

Equivalently, this means that the estimate $\hat{x}_t^{(i)}$ of each local observer (9) tends asymptotically to x_t for all state trajectories $x_t \in \mathcal{C}$ ensuing from the set

$$\mathcal{X} := \{x_0 \in \mathbb{R}^n : V(\mathbf{1}\otimes x_0) \leqslant \omega\}$$

(recall that $\hat{x}_0^{(i)} = 0$ for all $i \in \mathcal{N}$). Clearly, to guarantee asymptotic convergence of each $\hat{x}_t^{(i)}$ to x_t for all state trajectories x_t in $\mathcal{C}$ ensuing from $\mathcal{C}$ we must prove that $\mathcal{X}$ contains $\mathcal{C}$. We will take care of this in section 5.4.

Remark 5.1 *(Overlapping measurements). In the case of overlapping measurements among nodes (i.e. $y^{(j)}$ and $y^{(i)}$ for some $j \neq i$ have at least one component in common) both (27) and (28) are false. However, conclusions in (29) and (30) remain the same if in (29), (30) and (31) we change C into the matrix C_0 given by the $r < p$ independent rows $C_{j_1}, \ldots, C_{j_r}$ of C and the definition of R as $R = \Lambda C_0(\epsilon^{-\mathfrak{r}} \diamond G \diamond \epsilon^{-\mathfrak{r}})(\Lambda C_0)^T$, where Λ is a diagonal $r \times r$ matrix such that $\Lambda = \mathrm{diag}\{\#j_1, \ldots, \#j_r\}$ and $\#j_i$ is the square root of the number of repetitions of the measurement y_{j_i}, $i = 1, \ldots, r$, in the vector y.*

5.2 A centralized observer for (2) of the form (32)

As pointed out in the previous section, condition (25) is met once ϵ is selected, with K, P and R as in (31) and G as in (11), in such a way that:

$$\begin{aligned}
\mathcal{R}^{(g)}(x, \hat{e}^{(g)}) := 2\hat{e}^{(g)T}P\Big\{&(A - KC)\hat{e}^{(g)} \\
&+ \Delta(\phi\circ\sigma)\left(x, x - \hat{e}^{(g)}\right) \\
&- KC\Delta(C^T\psi\circ\sigma)\left(x, x - \hat{e}^{(g)}\right)\Big\} < 0
\end{aligned} \tag{37}$$

for all $x \in \mathcal{C}$ and $\hat{e}^{(g)} \neq 0$ ($\mathcal{R}^{(g)}$ is defined in (29) and here normalized by $\mathrm{Tr}\{D\} = N$). Here the collective observability assumption, i.e. (C,A) observable (and in the form (3)), comes into play. Define $\tilde{e}^{(g)} := (I - GA^T)\hat{e}^{(g)}$ and $\bar{c}, c$ and Γ be as in the previous section. As a matter of fact, for all $\epsilon \geqslant \epsilon_1$, ϵ_1 as in (15), using the definitions (11), the group properties of the dilations (sections B.1 and B.2) and lemma B.5, we obtain after some computations that for all $x \in \mathcal{C}$ and $\hat{e}^{(g)}$

$$\mathcal{R}^{(g)}(x, \hat{e}^{(g)}) \leqslant -\|\epsilon^{-\mathfrak{r}+\mathfrak{g}} \diamond \tilde{e}^{(g)}\|^2.$$

5.3 A centralized observer of the form (34) for the linearization of (33) around its trajectories

As pointed out in the section 5.1, condition (26) is met once ϵ is selected, with the same K, P and R and G used in section 5.2, in such a way that:

$$\mathcal{R}^{(L)}(x,\hat{e}^{(L)}) := 2\hat{e}^{(L)T}P\Big\{A-KC+\frac{\partial\phi}{\partial x}(x) - KCC^T\frac{\partial\psi}{\partial x}(x)\Big\}\hat{e}^{(L)} < 0 \tag{38}$$

for all $x \in \mathcal{C}$ and $\hat{e}^{(L)} \neq 0$ ($\mathcal{R}^{(L)}$ is defined in (30) and here normalized by $\mathrm{Tr}\{D\} = N$). Pick $\epsilon_2 \geqslant \epsilon_1$ such that for all $\epsilon \geqslant \epsilon_2$ and $x \in \mathcal{C}$

$$-\mathbf{1}_n c \preceq \epsilon^{-\mathfrak{r}} \diamond x \preceq \mathbf{1}_n c \tag{39}$$

and let $\epsilon \geqslant \epsilon_2$. Define $\widetilde{e}^{(L)} := (I - GA^T)\hat{e}^{(L)}$. As a matter of fact, by remark A.1 and on account, in particular, of inequality (A.1), we obtain after some computations that for all $x \in \mathcal{C}$ and $\hat{e}^{(L)}$

$$\mathcal{R}^{(L)}(x,\hat{e}^{(L)}) \leqslant -\|\epsilon^{-\mathfrak{r}+\mathfrak{g}} \diamond \widetilde{e}^{(L)}\|^2.$$

5.4 Enlarging the region of convergence for the estimation error

Our final task is to guarantee asymptotic convergence of each $\hat{x}_t^{(i)}$ to x_t for all state trajectories x_t in $\mathcal{C}$ ensuing form $\mathcal{C}$ and to do this, as discussed at the end of section 5.1, we must finally choose $\omega > 0$ and $\epsilon = \epsilon^* \geqslant \epsilon_2$ (see definition of ϵ_1 in (15) and ϵ_2 in (39)) in such a way that

$$\mathcal{X} := \{x_0 \in \mathbb{R}^n : V(\mathbf{1} \otimes x_0) \leqslant \omega\} \supset \mathcal{C}. \tag{40}$$

To this aim, we set $\overline{\mathfrak{g}} := \max_j \mathfrak{g}_j$, $\underline{\mathfrak{g}} := \min_j \mathfrak{g}_j$ and $\omega := \epsilon^{2(\overline{\mathfrak{g}}-\underline{\mathfrak{g}})}$. But

$$\begin{aligned} V(\mathbf{1} \otimes x_0) &= (\mathbf{1} \otimes x_0)^T (D \otimes P)(\mathbf{1} \otimes x_0) \\ &= N\|\epsilon^{-\mathfrak{r}} \diamond (I - GA^T)x_0\|^2 \end{aligned} \tag{41}$$

(recall the definition of P and that $\mathrm{Tr}(D) = N$). Using (iv) of lemma B.5 and (14), for all $\epsilon \geqslant \epsilon_2$ and $x_0 \in \mathcal{C}$

$$\|\epsilon^{-\mathfrak{r}} \diamond (I - GA^T)x_0\|^2 \leqslant \epsilon^{2(\overline{\mathfrak{g}}-\underline{\mathfrak{g}})}\|(I + \Gamma A^T)(\epsilon^{-\mathfrak{r}} \diamond x_0)\|^2.$$

Since there clearly exists $\epsilon^* \geqslant \epsilon_2$ such that for all $\epsilon \geqslant \epsilon^*$

$$\max_{x_0 \in \mathcal{C}} \|(I + \Gamma A^T)(\epsilon^{-\mathfrak{r}} \diamond x_0)\|^2 \leqslant \frac{1}{N} \tag{42}$$

then, as a consequence of (41), $\max_{x_0 \in \mathcal{C}} V(\mathbf{1} \otimes x_0) \leqslant \omega$ for all $\epsilon \geqslant \epsilon^*$. This condition with (41) gives (40).

Remark 5.2 *(Non zero initial conditions for the local observers). Each local observer (9) is initialized at $\hat{x}_0^{(i)} \neq 0$, $i = 1, \ldots, N$, for simplicity. If each local observer (9) is initialized at some nonzero $\hat{x}_0^{(i)} \in \mathcal{C}$, $i = 1, \ldots, N$, then (40) must be changed accordingly into*

$$\mathcal{X} := \{(x_0, \hat{x}_0) \in \mathbb{R}^n \times \mathbb{R}^{nN} : V(x_0 - \hat{x}_0) \leqslant \omega\} \supset \mathcal{C} \times \mathcal{C}^{\times N}.$$

where $\mathcal{C}^{\times N} := \mathcal{C} \times \cdots \times \mathcal{C}$ (N times) and $\hat{x}_0 := (\hat{x}_0^{(1)T}, \ldots, \hat{x}_0^{(N)T})^T$ is the vector of initial conditions for the distributed observer. Consistently, in (41)-(42) x_0 must be replaced with $x_0 - \hat{x}_0$.

6 Robustness

Once the design of the distributed observer has been clearly defined in the absence of system disturbances or uncertainties, we consider the more general model (8) where $\delta \in \mathbb{R}^s$ is a (time-varying) bounded disturbance, i.e. $\|\delta_t\| \leqslant \delta_\infty$ for all $t \geqslant 0$. We will design a distributed nonlinear observer for the system (8) over the given network $\mathcal{G}$. The local observer at node i (9) is modified as follows:

$$\begin{aligned} \dot{\hat{x}}_t^{(i)} &= A\hat{x}_t^{(i)} + \phi\left(\sigma_{c\epsilon^{\mathfrak{r}}}(\hat{x}_t^{(i)}), 0\right) \\ &\quad + \frac{\mathrm{Tr}(D)}{D_{i,i}} L^{(i)} Z_t^{(i)} + \gamma(\epsilon) P^{-1} K_t^{(i)},\ i \in \mathcal{N},\ \hat{x}_0^{(i)} = 0, \end{aligned} \tag{43}$$

with $\gamma \in \mathcal{K}_\infty$ and innovation process

$$Z_t^{(i)} := y_t^{(i)} - H^{(i)}\hat{x}_t^{(i)} - \psi^{(i)}\left(\sigma_{c\epsilon^{\mathfrak{r}}}(\hat{x}_t^{(i)}), 0\right), \tag{44}$$

consensus term $K_t^{(i)}$ as in (10) and matrices $L^{(i)}, P, R^{(i)}$ and G as in (11). The parameters γ, c, ϵ and Γ will depend on the system nonlinearities $\phi(x, 0)$ and $\psi(x, 0)$ which contain much less information than $\phi(x, \delta)$ and $\psi(x, \delta)$ and guarantee the robustness of the distributed observer against disturbance/uncertainties δ. The task of a distributed observer for (2) is to cooperatively estimate the state of the system (2) in the presence of the disturbance/uncertainty δ. For a formulation of our problem we will refer to the following robust notion of omniscience.

Definition 6.1 *(Asymptotic Omniscience (relatively to $\mathcal{C} \subset \mathbb{R}^n$ with error tolerance $\zeta \in \mathbb{R}_>$)): A distributed observer $\{\hat{x}^{(i)}\}_{i \in \mathcal{N}}$ achieves asymptotic omniscience for (2) if for all the state trajectories x_t of (8) in $\mathcal{C} \subset \mathbb{R}^n$ we have $\limsup_{t \to +\infty} \|\hat{x}_t^{(i)} - x_t\| \leqslant \zeta$ for all $i \in \mathcal{N}$, i.e. the state estimate $\hat{x}_t^{(i)}$ at each node i asymptotically converges to x_t with a maximum tolerated error ζ.*

Our assumptions on the system (8) must be strengthened as follows for the presence of the additional variable δ in ϕ and ψ

(**H0**)′(*state boundedness*): *the state trajectories of (8) are contained in some known compact set* $\mathcal{C} \subset \mathbb{R}^n$,

(**H1**)′ (*incremental homogeneity in the upper bound*): *for some degrees* $\mathfrak{g}, \hat{\mathfrak{g}}, \mathfrak{h} \in \mathbb{R}^n$ *and weights* $\mathfrak{r}, \mathfrak{s} \in \mathbb{R}^n_>$ *satisfying* $\mathfrak{g}_{j+1} < \hat{\mathfrak{g}}_j \leqslant \mathfrak{g}_j$ *for* $j \notin S$, *where* $\hat{\mathfrak{g}} \in \mathbb{R}^n$ *is defined in (13), and*

$$\max_j\{\mathfrak{h}_j - \mathfrak{s}_j\} < \min_j\{\mathfrak{g}_j\} - \max_j\{\mathfrak{r}_j\}, \tag{45}$$

(i) $\phi(x,\delta)$ *is incrementally homogeneous in the upper bound with quadruples*

$$\left\{\begin{pmatrix}\mathfrak{r}\\ \mathfrak{s}\end{pmatrix}, \mathfrak{r}+\hat{\mathfrak{g}}, \begin{pmatrix}\mathfrak{g}\\ \mathfrak{h}\end{pmatrix}, \Phi\left(\begin{pmatrix}x'\\ x''\end{pmatrix}, \begin{pmatrix}\delta'\\ \delta''\end{pmatrix}\right)\right\},$$

with $\Phi(0,0) = BF$, B *as in (7) and for some* $F \in \mathbb{R}^{p\times n}$,

(ii) $C^T\psi(x,\delta)$ *is incrementally homogeneous in the upper bound with quadruples*

$$\left\{\begin{pmatrix}\mathfrak{r}\\ \mathfrak{s}\end{pmatrix}, \mathfrak{r}-\mathfrak{g}, \begin{pmatrix}\mathfrak{g}\\ \mathfrak{h}\end{pmatrix}, C^T\Psi\left(\begin{pmatrix}x'\\ x''\end{pmatrix}, \begin{pmatrix}\delta'\\ \delta''\end{pmatrix}\right)\right\},$$

with $\Psi(0,0) = \alpha C$ *and* $\alpha \in [0,1)$,

(**H2**)′ (*uniform incremental homogeneity in the* 0*-limit*): $\phi(x,0)$ *is uniformly incrementally homogeneous in the* 0*-limit with triple* $(\mathbf{1}_n, \mathbf{1}_n, \frac{\partial\phi}{\partial x'}(x',0)x'')$ *and* $C^T\psi$ *is uniformly incrementally homogeneous in the* 0*-limit with triple* $(\mathbf{1}_n, \mathbf{1}_n, C^T\frac{\partial\psi}{\partial x'}(x',0)x'')$.

For example $\phi(x,\delta) = (0, -x_1 + x_2(1-x_1^2) + \delta)^T$ satisfies (**H1**)′ with $\mathfrak{s} = \mathfrak{r}_1 = 1/2$, $\mathfrak{r}_2 = 2$, $\mathfrak{g}_1 = 1$, $\mathfrak{g}_2 = 1/2$ and $\mathfrak{h} = -2$, $B = (0,1)^T$ and $F = (1,1,1)$. The main result of this section concerns the achievement of omniscience in the above specified sense and it is stated as follows.

Theorem 6.1 *Consider a network* $\mathcal{N}$ *described by a directed strongly connected graph* $\mathcal{G}$ *and a system (8) with* (C,A) *observable (collective observability). Given a compact set* $\mathcal{C} \subset \mathbb{R}^n$, $\zeta \in \mathbb{R}_>$, *and under assumptions (***H0***)′-(***H2***)′ there exist* $\epsilon^* \geqslant 1$, *diagonal positive definite* Γ *and* D, $\gamma \in \mathcal{K}_\infty$ *and* $c \in \mathbb{R}_>$ *such that for all* $\epsilon \geqslant \epsilon^*$ *the distributed observer* $\{\hat{x}^{(i)}\}_{i\in\mathcal{N}}$ *over the given network* $\mathcal{N}$ *described by (43) achieves asymptotic omniscience for (2) relatively to* $\mathcal{C}$ *with error tolerance* ζ.

(*Sketch of the proof*). The proof closely follows the proof of theorem 4.1. With $\epsilon > 1$, $\underline{\mathfrak{g}} := \min_j \mathfrak{g}_j$ and $\overline{\mathfrak{h}-\mathfrak{s}} := \max_j\{\mathfrak{h}_j - \mathfrak{s}_j\}$, using the assumptions (**H0**)′-(**H2**)′ we obtain for a given $\omega > 0$ some $\Pi, \beta > 0$ such that

$$\begin{aligned}&\forall e : \frac{4\Pi N}{\beta}\epsilon^{-2(\underline{\mathfrak{g}}-\overline{\mathfrak{h}-\mathfrak{s}})} \leqslant V(e) \leqslant \omega\\ &\Rightarrow \dot{V} \leqslant -\frac{\beta}{4N}\epsilon^{2\underline{\mathfrak{g}}}V(e),\ \forall x \in \mathcal{C}.\end{aligned} \tag{46}$$

This means that the estimation error e_t enters (in finite time) the set

$$\mathcal{B} := \{e \in \mathbb{R}^{nN} : V(e) \leqslant \frac{4\Pi N}{\beta}\epsilon^{-2(\underline{\mathfrak{g}}-\overline{\mathfrak{h}-\mathfrak{s}})}\}$$

and remains thereinafter, for all state trajectories $x_t \in \mathcal{C}$ ensuing from the set $\mathcal{X}$ in (40). On account of (45) in assumption (**H1**)′, we can select $\epsilon > 1$ in such a way that $\mathcal{C} \subset \mathcal{X}$ and at the same time $\mathcal{B} \subset \{e \in \mathbb{R}^n : \|e\| \leqslant \zeta\}$, which proves omniscience with error tolerance ζ.

7 Switching topologies

Once the design of the distributed observer has been clearly defined for networks with time-invariant topologies, we consider the distributed state estimation problem for (8) over a network $\mathcal{G}_t$ with switching topology. We assume that the directed graph $\mathcal{G}_t$ is described by an adjacency matrix $\mathcal{A}_t$ which switches at (increasing) times $t = t_j$, $j \in \mathbb{N}$, with $\inf_{j\in\mathbb{N}}(t_{j+1} - t_j) > 0$ (*dwell time*), among a collection of fixed topologies $\{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$, l a given integer. We also assume that for all the possible topologies $\{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$ the graph remains strongly connected. Accordingly, the graph has an associated Laplacian $\mathcal{L}_t$ which switches at times $t = t_j$, $j \in \mathbb{N}$, among a finite collection $\{\mathcal{L}_1, \ldots, \mathcal{L}_l\}$. By the strong connectivity of the graph under switching, there exist a matrix D_t, diagonal and positive definite at each $t \geqslant 0$, and T_t, nonsingular at each $t \geqslant 0$, such that $D_{t_j}\mathcal{L}_{t_j} + \mathcal{L}_{t_j}^T D_{t_j}$ is positive semi-definite and $T_{t_j}^T(D_{t_j}\mathcal{L}_{t_j} + \mathcal{L}_{t_j}^T D_{t_j})T_{t_j} = \{0, \lambda_{2,t_j}, \ldots, \lambda_{N,t_j}\}$ (see lemma B.1) for each $j \in \mathbb{N}$ with $0 < \lambda_{i,t_j} \leqslant \lambda_{i+1,t_j}$, $i = 2, \ldots, N-1$.

With this in mind, the local observer at node i (43) is modified as follows: for $t \in [t_j, t_{j+1})$

$$\begin{aligned}\dot{\hat{x}}_t^{(i)} =& A\hat{x}_t^{(i)} + \phi\left(\sigma_{c\epsilon^{\mathfrak{r}}}(\hat{x}_t^{(i)}), 0\right)\\ &+ \frac{\mathrm{Tr}(D_t)}{[D_t]_{i,i}}L_t^{(i)}Z^{(i)} + \gamma(\epsilon)P^{-1}K_t^{(i)},\ i \in \mathcal{N},\ \hat{x}_0^{(i)} = 0,\end{aligned} \tag{47}$$

with $\gamma \in \mathcal{K}_\infty$, innovation process $Z_t^{(i)}$ as in (44), consensus term $K_t^{(i)} := \sum_{j=1}^N [\mathcal{A}_t]_{i,j}(\hat{x}^{(j)} - \hat{x}^{(i)})$, where $[\mathcal{A}_t]_{i,j}$ is the (i,j) entry of the adjacency matrix $\mathcal{A}_t$ of the given network $\mathcal{G}_t$, and matrices $L^{(i)}, P, R^{(i)}$ and G as in (11). Clearly, the switching times $\{t_j\}_{j\in\mathbb{N}}$ are known at each sensing node.

The main result of this section concerns the achievement of omniscience in the usual sense and it is stated with no proof as follows.

Theorem 7.1 *Consider a network* $\mathcal{N}_t$*described by a graph* $\mathcal{G}_t$ *with switching topology* $\mathcal{A}_t \in \{\mathcal{A}_1, \ldots, \mathcal{A}_l\}$, l *a given integer, and a system (8) with* (C,A) *observable*

(collective observability). Given a compact set $\mathcal{C} \subset \mathbb{R}^n$ *and* $\zeta \in \mathbb{R}_>$*, under assumptions* (**H0**)′-(**H2**)′ *there exist* $\epsilon^* \geqslant 1$, $t^* > 0$, *diagonal positive definite* Γ, D_t *diagonal and positive definite at each* $t \geqslant 0$, $\gamma \in \mathcal{K}_\infty$ *and* $c \in \mathbb{R}_>$ *such that for all* $\epsilon \geqslant \epsilon^*$ *the distributed observer* $\{\hat{x}^{(i)}\}_{i\in\mathcal{N}}$ *over the given network described by (47) achieves asymptotic omniscience for (2) relatively to* $\mathcal{C}$ *with error tolerance* ζ *as long as* $\inf_{j\in\mathbb{N}}(t_{j+1} - t_j) \geqslant t^*$.

Therefore, asymptotic omniscience is obtained at the price of a sufficiently large dwell time.

8 Example and simulations

To test our distributed observer we consider a simple plant built on coupled Van der Pol oscillators (VPO: [7]). The plant has the form (2) with:

$$A = I_N \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; C = I_N \otimes (1 \quad 0), \; N \geqslant 2,$$

$$\phi(x) = \text{diag}\left\{ \begin{pmatrix} 0 \\ f_1(x) \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ f_N(x) \end{pmatrix} \right\}$$

with $x = (x_1^T, \ldots, x_N^T)^T$, $x_i \in \mathbb{R}^2$, $i = 1, \ldots, N$, and

$$\begin{aligned} f_i(x) = &-\mu x_{i,1} + \mu(1 - x_{i,1}^2)x_{i,2} \\ &+ a(x_{1,1} + \ldots + x_{i-1,1} - (N-1)x_{i,1}) \\ &+ b(x_{1,2} + \ldots + x_{i-1,2} - (N-1)x_{i,2}), \end{aligned}$$

parameters $\mu = 5$ and coupling coefficients $a = 2$ and $b = 0.1$. We have chosen $N = 12$ (12 nodes in the graph) so that the state space has dimension $n = 2N = 24$. As compact set $\mathcal{C}$ which contains the system's state

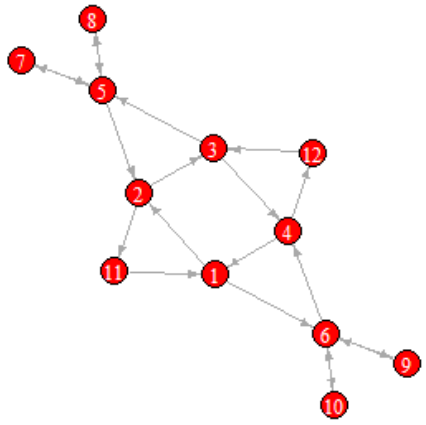

Fig. 1. communication directed graph

trajectories, we have taken $\mathcal{C} = \{x \in \mathbb{R}^{2n} : \quad \|x\| \leqslant 15\}$ (the system has a limit cycle contained in this region). Assumptions (**H1**)-(**H2**) are satisfied with $\mathfrak{r}_{2j+1} = 1$, $\mathfrak{r}_{2(j+1)} = 3$, $\mathfrak{g}_j = 1$, $j \in \mathbb{N}$ (in this example $S = \{2j : j \in \mathbb{N}\}$) and

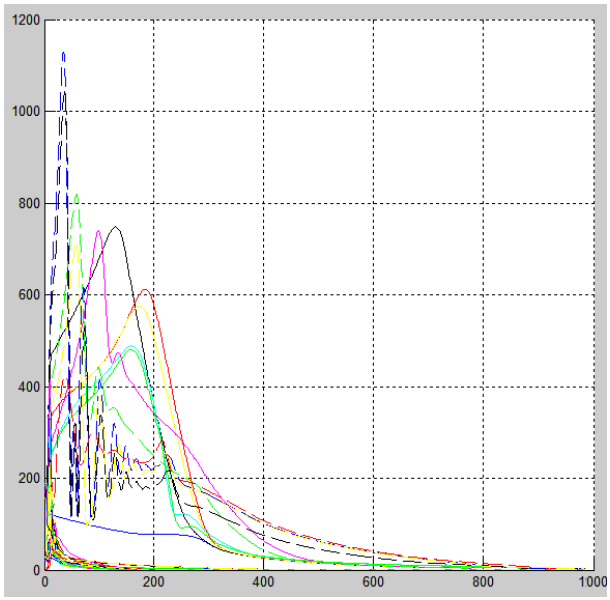

Fig. 2. distributed estimation error: omniscience over time

$$B = \text{diag}\{(0,1)^T, (0,1)^T, \cdots, (0,1)^T, (0,1)^T\}$$

$$F = \begin{pmatrix} \mu_a & \mu_b & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ a & b & \mu_a & \mu_b & & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & & \\ a & b & a & b & \cdots & \mu_a & \mu_b & 0 & 0 \\ a & b & a & b & \cdots & a & b & \mu_a & \mu_b \end{pmatrix}$$

with $\mu_a := \mu + a(N-1)$ and $\mu_b := \mu + b(N-1)$. We consider the strongly connected directed graph in Fig. 1 and each vertex measuring only the position of a single VPO. Following the procedure in IV.D we have chosen $\epsilon = 11.00$ and $\gamma = 12.00$ and designed our distributed observer (9). In Fig. 2 we plot the estimation error of each local observer.

9 Conclusions

In this paper we have proposed a novel distributed observer design for a general class of nonlinear systems. The information among the local observers is exchanged over a strongly connected and directed digraph. The local observer attemps to estimate the entire state of the plant and the asymptotic omniscience is studied and proved via a Lyapunov approach. Theoretical results have been confirmed by the numerical simulation. We also discussed robustness and switching topologies issues. In future research, we plan to focus on global distributed nonlinear scenarios.

A Incremental homogeneity

The notion of (incremental) homogeneity in a generalized sense has been introduced in [2] in the context of semi-global stabilization and observer design problems. Here we recall this notion in a slightly more general form. Let $\Delta f(x', x'') := f(x') - f(x'')$ and if f is the identity function we simply write $\Delta(x', x'') := x' - x''$.

Definition A.1 *A parametrized function $\phi_\epsilon \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^l)$, $\epsilon > 0$, is said to be incrementally homogeneous (i.h.) with quadruple $\{\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi(x', x'')\}$ if there exist $\mathfrak{d} \in \mathbb{R}^l, \mathfrak{h} \in \mathbb{R}^n$, $\mathfrak{r} \in \mathbb{R}^n_>$ and $\Phi \in \mathbf{C}^0(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^{l\times n})$ such that for all $\epsilon > 0$ and $x', x'' \in \mathbb{R}^n$*

$$\Delta\phi_\epsilon(\epsilon^{\mathfrak{r}} \diamond x', \epsilon^{\mathfrak{r}} \diamond x'') = \epsilon^{\mathfrak{d}} \diamond \Phi(x', x'')\Delta(\epsilon^{\mathfrak{h}} \diamond x', \epsilon^{\mathfrak{h}} \diamond x'')$$

When $x'' := 0$, we say that ϕ_ϵ *is homogeneous with quadruple* $\{\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi'(x)\}$ where $\Phi'(x) = \Phi(x, 0)$.

Example A.1 *The function $\phi_\epsilon(x) := x_1 + x_2^3$ (in this case ϕ_ϵ does not depend on the dilating parameter) is i.h. with quadruple $\{\mathfrak{r}, 0, \mathfrak{h}, \Phi(x', x'')\}$, where $\mathfrak{r} := (1, 2)^T$, $\mathfrak{h} := (1, 6)^T$ and $\Phi(x', x'') := \big(1, (x_2')^2 + (x_2'')^2 + x_2''x_2'\big)$.*

There are functions, like $\sin x$, which are not i.h. but behaves in the upper bound as an i.h. function. This motivates the following definition ($\langle\!\langle a \rangle\!\rangle$ denotes the column vector of the absolute values of the elements of $a \in \mathbb{R}^n$).

Definition A.2 *A parametrized function $\phi_\epsilon \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^l)$, $\epsilon > 0$, is said to be incrementally homogeneous in the upper bound (i.h.u.b.) with quadruple $\{\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi(x', x'')\}$ if there exist $\mathfrak{d} \in \mathbb{R}^l, \mathfrak{h} \in \mathbb{R}^n$, $\mathfrak{r} \in \mathbb{R}^n_>$, $\Phi \in \mathbf{C}^0(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^{l\times n}_\geqslant)$ such that for all $\epsilon \geqslant 1$ and $x', x'' \in \mathbb{R}^n$*

$$\langle\!\langle \Delta\phi_\epsilon(\epsilon^{\mathfrak{r}} \diamond x', \epsilon^{\mathfrak{r}} \diamond x'') \rangle\!\rangle \preceq \epsilon^{\mathfrak{d}} \diamond \Phi(x', x'') \Big\langle\!\Big\langle \Delta(\epsilon^{\mathfrak{h}} \diamond x', \epsilon^{\mathfrak{h}} \diamond x'') \Big\rangle\!\Big\rangle$$

When $x'' := 0$, we say that ϕ_ϵ *is homogeneous in the upper bound with quadruple* $\{\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi'(x)\}$ where $\Phi'(x) = \Phi(x, 0)$.

Example A.2 *The function $\phi_\epsilon(x) := \epsilon \begin{pmatrix} x_2 & x_2^3 g(x_1) \end{pmatrix}^T$, $g \in \mathbf{C}^0(\mathbb{R}, \mathbb{R})$ any bounded and globally Lipschitz function, is i.h.u.b. with triple $\{\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi(x', x'')\}$, where $\mathfrak{r} := (1, 2)^T$, $\mathfrak{d} := (3, 7)^T$, $\mathfrak{h} := (1, 0)^T$ and $\Phi(x', x'')$ defined as follows:*

$$\begin{aligned}
&[\Phi(x', x'')]_{1,1} = 0,\ [\Phi(x', x'')]_{1,2} = 1\\
&[\Phi(x', x'')]_{2,1} = (x_2'')^3 \frac{|g(x_1') - g(x_1'')|}{|x_1' - x_1''|},\\
&[\Phi(x', x'')]_{2,2} = |(x_2')^2 + (x_2'')^2 + x_2'x_2''||g(x_1')|.
\end{aligned}$$

Remark A.1 *An important consequence of the definitions A.1 and A.2 is the following. If $\phi_\epsilon \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^l)$, $\epsilon > 0$, is incrementally homogeneous in the upper bound with quadruple $\{\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi(x', x'')\}$ then by letting*

$$x^i(\delta) = (x_1, \ldots, x_{i-1}, x_i + \delta, x_{i+1}, \ldots, x_n)$$

and $\phi_{\epsilon,j}$ being the j-th component of ϕ_ϵ and $\Phi_{j,i}$ the (j, i)-th entry of Φ, we have for $\epsilon \geqslant 1$

$$|\Delta\phi_{\epsilon,j}(\epsilon^{\mathfrak{r}} \diamond x, \epsilon^{\mathfrak{r}} \diamond x^i(\delta))| \leqslant \epsilon^{\mathfrak{d}_j + \mathfrak{h}_i} \diamond \Phi_{j,i}(x, x^i(\delta))|\delta|.$$

If the partial derivatives of ϕ exist and are continuous, by dividing both part of the above inequality by $\epsilon^{\mathfrak{r}_i}|\delta|$ and letting $\delta \to 0$ we get for $\epsilon \geqslant 1$

$$\begin{aligned}
\left|\frac{\partial\phi_{\epsilon,j}}{\partial x_i}(\epsilon^{\mathfrak{r}} \diamond x)\right| &= \lim_{\delta\to 0} \frac{|\Delta\phi_{\epsilon,j}(\epsilon^{\mathfrak{r}} \diamond x, \epsilon^{\mathfrak{r}} \diamond x^i(\delta))|}{|\epsilon^{\mathfrak{r}_i}\delta|}\\
&\leqslant \epsilon^{\mathfrak{d}_j + \mathfrak{h}_i - \mathfrak{r}_i} \diamond \Phi_{j,i}(x, x).
\end{aligned}$$

As a consequence, if we consider the function $\eta_\epsilon : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$: $\eta_\epsilon(x', x'') := \frac{\partial\phi_\epsilon}{\partial x'}(x')x''$ we obtain the following inequality for $\epsilon \geqslant 1$ and $x', x'' \in \mathbb{R}^n$

$$\langle\!\langle \eta_\epsilon(\epsilon^{\mathfrak{r}} \diamond x', \epsilon^{\mathfrak{r}} \diamond x'') \rangle\!\rangle \preceq \epsilon^{\mathfrak{d}} \diamond \Phi(x', x') \big\langle\!\big\langle \epsilon^{\mathfrak{h}} \diamond x'' \big\rangle\!\big\rangle \quad \text{(A.1)}$$

i.e. η_ϵ is homogeneous in the upper bound with quadruple $\left\{ \begin{pmatrix} \mathfrak{r} \\ \mathfrak{r} \end{pmatrix}, \mathfrak{d}, \begin{pmatrix} 0 \\ \mathfrak{h} \end{pmatrix}, \begin{pmatrix} 0_{n\times n} & \Phi(x', x'') \end{pmatrix} \right\}$.

Properties of incremental homogeneity can be found in [2]. Incremental homogeneity in the upper bound as introduced above cope with large values of the dilating parameter (i.e. $\epsilon >> 1$). Another related important incremental homogeneity notion for the behavior of a function near 0 (i.e. $\epsilon << 1$) is the following.

Definition A.3 *A function $\phi \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R})$ is said to be uniformly incrementally homogeneous in the 0-limit (u.i.h-0) with triple $\{\mathfrak{r}, \mathfrak{d}, \Phi_0(x', x'')\}$, weights $\mathfrak{r}$, degree $\mathfrak{d}$ and limit function $\Phi_0(x', x'')$, if there exist $\mathfrak{d} \in \mathbb{R}$, $\mathfrak{r} \in \mathbb{R}^n_>$ and $\Phi_0 \in \mathbf{C}^0(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ such that for all compact sets $\mathcal{C} \subset \mathbb{R}^n$ and $\mathcal{S} \subset \mathbb{R}^n$*

$$\lim_{\lambda\to 0} \max_{\substack{x'\in\mathcal{C}\\ x''\in\mathcal{S}}} \left|\frac{\Delta\phi(x', x' + \lambda^{\mathfrak{r}} \diamond x'')}{\lambda^{\mathfrak{d}}} - \Phi_0(x', x'')\right| = 0.$$

A function $\phi \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^l)$ is said to be uniformly incrementally homogeneous in the 0-limit with triple $\{\mathfrak{r}, \mathfrak{d}, \Phi_0(x', x'')\}$, weights $\mathfrak{r}$, degrees $\mathfrak{d}$ and limit function $\Phi_0(x', x'')$, if ϕ_j, $j = 1, \ldots, l$, is uniformly incrementally homogeneous in the 0-limit with triple $\{\mathfrak{r}, \mathfrak{d}_j, \Phi_{0,j}(x', x'')\}$ and $\Phi_{0,j}(x', x'')$ is the j-th row of $\Phi_0(x', x'')$.

In this case the limit function $\Phi_0(x', x'')$ is the stack of the rows $\Phi_{0,j}(x', x'')$, $j = 1, \ldots, l$, i.e. $\Phi_0(x', x'') =$

$$(\Phi_{0,1}(x', x'') \quad \cdots \quad \Phi_{0,l}(x', x''))^T.$$

Example A.3 *The function* $\phi(x) := x + x^3$ *is u.i.h-*0 *with triple* $\{1, 1, x'' + 3(x')^2 x''\}$.

B Auxiliary results and properties

B.1 Properties of block matrices (C, A)

For any diagonal $G \in \mathbb{R}^{nN \times nN}$ the matrices A and C in (3) have the following properties summed below without further comments:

$$CA^T = 0,\ CC^T = I, \tag{B.1}$$

$$(GA^T)^j = 0,\ \forall j \geqslant n,\ (I - GA^T)^{-1} = \sum_{j=0}^{n-1} (GA^T)^j \tag{B.2}$$

$$CGA^T = 0,\ C(I - GA^T)^{-1} = C, \tag{B.3}$$

$$GA^T A = A^T AGA^T A,\ GAA^T = AA^T GAA^T, \tag{B.4}$$

$$C^T C = I - A^T A,\ C^T CA^T A = 0, \tag{B.5}$$

$$A^T A(I - A^T A) = 0,\ AA^T (I - AA^T) = 0. \tag{B.6}$$

B.2 Group properties of dilations and its action

For any diagonal matrix D: $D \diamond \epsilon^{\mathfrak{r}} = \epsilon^{\mathfrak{r}} \diamond D$ (*commutativity*). For any matrices R, S with suitable dimensions:

$$(RS) \diamond \epsilon^{\mathfrak{r}} = R(S \diamond \epsilon^{\mathfrak{r}}),\ \epsilon^{\mathfrak{r}} \diamond (RS) = (\epsilon^{\mathfrak{r}} \diamond R)S$$
$$(R \diamond \epsilon^{\mathfrak{r}})S = R(\epsilon^{\mathfrak{r}} \diamond S),\ S(\epsilon^{\mathfrak{r}} \diamond R) = (S \diamond \epsilon^{\mathfrak{r}})R$$

(*associativity*) and $(R \diamond \epsilon^{\mathfrak{r}})^T = \epsilon^{\mathfrak{r}} \diamond R^T$.

B.3 Auxiliary lemmas for graphs

First, we borrow two lemmas from [14], [17], [20] and [5].

Lemma B.1 *Assume* $\mathcal{G}$ *is a strongly connected directed graph with Laplacian* $\mathcal{L}$. *There exists a unique positive row vector* $d = (d_1, \ldots, d_N)$ *such that* $d\mathcal{L} = 0$ *and* $d\mathbf{1}_N = N$. *If* $D := \mathrm{diag}\{d_1, \ldots, d_N\}$ *then* $\hat{\mathcal{L}} := D\mathcal{L} + \mathcal{L}^T D$ *is positive semi-definite and* $\hat{\mathcal{L}}\mathbf{1}_N = 0$.

Let D and $\hat{\mathcal{L}}$ be as in the above lemma.

Lemma B.2 *For a strongly connected directed graph* $\mathcal{G}$ *with Laplacian* $\mathcal{L}$, 0 *is a simple eigenvalue of* $\hat{\mathcal{L}}$. *Furthermore, its eigenvalues can be ordered as* $\lambda_1 = 0 < \lambda_2 \leqslant \lambda_3 \leqslant \cdots \leqslant \lambda_N$ *and there exists a* $N \times N$ *orthogonal matrix such that its first row is* $\frac{1}{\sqrt{N}}\mathbf{1}_N^T$ *and*

$$T(D\mathcal{L} + \mathcal{L}^T D)T^T = \Lambda := \mathrm{diag}\{0, \lambda_2, \ldots, \lambda_N\}. \tag{B.7}$$

B.4 Auxiliary lemmas for negative definite nonlinear forms

A crucial lemma in our nonlinear distributed estimation problem is the following.

Lemma B.3 *Let* $\mathcal{C} \subset \mathbb{R}^n$ *and* $\Omega \subset \mathbb{R}^{\times N} := \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ *(*N *times) be compact sets, both containing the origin, and a continuous function* $\mathcal{V} : \mathcal{C} \times \mathbb{R}^{\times N} \to \mathbb{R}$:

$$\mathcal{V}(x, e) := -\kappa(\gamma) e^T \Lambda e + \mathcal{W}(x, e)$$

be given for some $\kappa \in \mathcal{K}_\infty$ *and positive semidefinite* $\Lambda \in \mathbb{R}^{nN \times nN}$ *with continuous* $\mathcal{W} : \mathcal{C} \times \mathbb{R}^{\times N} \to \mathbb{R}$. *Moreover,* $\mathcal{W}(x, 0) = 0$ *for all* $x \in \mathcal{C}$. *If (i) there exists a continuous* $\mathcal{W}_0 : \mathcal{C} \times \mathbb{R}^{\times N} \to \mathbb{R}$ *such that for any* N*-tuple of compact sets* $\mathcal{S}^{(1)}, \ldots, \mathcal{S}^{(N)} \subset \mathbb{R}^n$ *with* $\mathcal{S}^{\times N} := \mathcal{S}^{(1)} \times \cdots \times \mathcal{S}^{(N)}$:

$$\lim_{\lambda \to 0} \max_{\substack{x \in \mathcal{C} \\ e \in \mathcal{S}^{\times N}}} \left| \frac{\mathcal{W}(x, \lambda e)}{\lambda^2} - \mathcal{W}_0(x, e) \right| = 0, \tag{B.8}$$

(ii) for each $e \in \Omega \backslash \{0\}$

$$e^T \Lambda e = 0 \Rightarrow \mathcal{W}(x, e) < 0,\ \forall x \in \mathcal{C}, \tag{B.9}$$

$$e^T \Lambda e = 0 \Rightarrow \mathcal{W}_0(x, e) < 0,\ \forall x \in \mathcal{C}, \tag{B.10}$$

there exists $\overline{\gamma} > 0$ *such that*

$$\mathcal{V}(x, e) < 0, \forall \gamma \geqslant \overline{\gamma},\ e \in \Omega \backslash \{0\},\ x \in \mathcal{C}.$$

PROOF. I) Using (i) we can select $\gamma^* > 0$ and an open ball $\mathcal{B} \subset \Omega$ around $e = 0$ such that

$$\mathcal{V}(x, e) < 0,\ \forall \gamma \geqslant \gamma^*,\ e \in \mathcal{B} \backslash \{0\},\ x \in \mathcal{C}. \tag{B.11}$$

Indeed, first notice that for each $x \in \mathcal{C}$, $e \in \mathbb{R}^{\times N}$ and $\mu > 0$

$$\begin{aligned} \frac{\mathcal{W}_0(x, \mu e)}{\mu^2} &= \frac{1}{\mu^2} \lim_{\lambda \to 0} \frac{\mathcal{W}(x, \lambda(\mu e))}{\lambda^2} \\ &= \lim_{\lambda \to 0} \frac{\mathcal{W}(x, (\lambda\mu)e))}{(\lambda\mu)^2} = \mathcal{W}_0(x, e) \end{aligned} \tag{B.12}$$

Let $\mathcal{S}^\lambda \subset \mathbb{R}^{\times N}$ be the sphere with radius λ centered at 0. Clearly, $\mathcal{S}^1$ is contained in a compact set of the form $\mathcal{S}^{\times N}$. We first show, by contradiction, that there exists a real number $\gamma_0 > 0$ satisfying:

$$-\kappa(\gamma) e^T \Lambda e + \mathcal{W}_0(x, e) < 0,\ \forall \gamma \geqslant \gamma_0,\ x \in \mathcal{C},\ e \in \mathcal{S}^1.$$

Suppose there is no such γ_0. This means there is a sequence $\{(x_i, e_i)\}_{i \in \mathbb{N}} \in \mathcal{C} \times \mathcal{S}^1$ which satisfies :

$$-ie_i^T \Lambda e_i + \mathcal{W}_0(x_i, e_i) \geqslant 0\ \forall i \in \mathbb{N}. \tag{B.13}$$

Since $\mathcal{C}\times\mathcal{S}^1$ is compact there exists a convergent subsequence $\{(x_{i_l},e_{i_l})\}_{l\in\mathbb{N}}\in\mathcal{C}\times\mathcal{S}^1$ to a point $\{(x_\infty,e_\infty)\}\in\mathcal{C}\times\mathcal{S}^1$. Since the functions $\eta:\mathbb{R}^{\times N}\to\mathbb{R}$, $e\mapsto\eta(e)=e^T\Lambda e$ and $\mathcal{W}_0$ are bounded on $\mathcal{S}^1$ and $\mathcal{C}\times\mathcal{S}^1$, respectively, and η is nonnegative, on account of (B.13) η must tend to zero as i_l tends to infinity. By continuity of η and $\mathcal{W}_0$ it follows that $\eta(e_\infty)=0$ and, on account of (B.13), $\mathcal{W}_0(x_\infty,e_\infty)\geqslant 0$. But this is a contradiction from (B.10). We conclude that there exist $\delta_0,\gamma_0>0$ such that

$$-\kappa(\gamma)\eta(e)+\mathcal{W}_0(x,e)\leqslant-\delta_0<0,\quad\forall\gamma\geqslant\gamma_0,\ x\in\mathcal{C},\ e\in\mathcal{S}^1. \tag{B.14}$$

On account of (B.8) and since η is a quadratic function, there exists $\lambda_0>0$ such that for all $\lambda\in(0,\lambda_0]$, $x\in\mathcal{C}$ and $e\in\mathcal{S}^1$

$$\mathcal{W}(x,\lambda e)\leqslant\lambda^2\mathcal{W}_0(x,e)+\lambda^2\frac{\delta_0}{2},\ \eta(\lambda e)=\lambda^2\eta(e), \tag{B.15}$$

which together give

$$-\kappa(\gamma_0)\eta(e)+\mathcal{W}(x,e)\leqslant-\lambda^2\kappa(\gamma_0)\eta(e)+\lambda^2\mathcal{W}_0(x,e)+\lambda^2\frac{\delta_0}{2}.$$

Using (B.14), we obtain that for all $\lambda\in(0,\lambda_0]$, $x\in\mathcal{C}$ and $e\in\mathcal{S}^1$: $-\kappa(\gamma_0)\eta(\lambda e)+\mathcal{W}(x,\lambda e)\leqslant-\lambda^2\frac{\delta_0}{2}$ and, since η is nonnegative, for all $\lambda\in(0,\lambda_0]$, $\gamma\geqslant\gamma_0$, $x\in\mathcal{C}$ and $e\in\mathcal{S}^1$: $-\kappa(\gamma)\eta(\lambda e)+\mathcal{W}(x,\lambda e)\leqslant-\lambda^2\frac{\delta_0}{2}$. This implies (B.11) with $\gamma^*=\gamma_0$ and $\mathcal{B}=\lambda_0\mathrm{int}(\mathcal{S}^1)=\mathrm{int}(\mathcal{S}^{\lambda_0})$ (int denotes the interior of a set).

II) To complete the proof, we will show that there exists $\gamma^{**}\geqslant\gamma^*$ such that $\mathcal{V}(x,e)<0$ for all $\gamma\geqslant\gamma^{**}$, for all $e\in\Omega\backslash\mathcal{B}$ and $x\in\mathcal{C}$ and our lemma will be proved with $\overline{\gamma}=\gamma^{**}$. Assume our claim be false. For each n there exists a point $(x_n,e_n)\in\mathcal{C}\times(\Omega\backslash\mathcal{B})$ such that $-\kappa(n)e_n^T\Lambda e_n+\mathcal{W}(x_n,e_n)\geqslant 0$. Since $\kappa\in\mathcal{K}_\infty$, for each m and for all $n\geqslant m$

$$-\kappa(m)e_n^T\Lambda e_n+\mathcal{W}(x_n,e_n)\geqslant 0.$$

But $\mathcal{C}\times(\Omega\backslash\mathcal{B})$ is compact in $\mathbb{R}^n\times\mathbb{R}^{\times N}$ so that there exists a point $(x^*,e^*)\in\mathcal{C}\times(\Omega\backslash\mathcal{B})$ to which the sequence $\{(x_n,e_n)\}_{n\in\mathbb{N}}$ converges as $n\to+\infty$. Therefore, using continuity of $\mathcal{W}$

$$-\kappa(m)(e^*)^T\Lambda e^*+\mathcal{W}(x^*,e^*)\geqslant 0 \tag{B.16}$$

for all $m>0$.

If $e^*\in(\Omega\backslash\mathcal{B})$ is such that $(e^*)^T\Lambda e^*=0$, since $\{0\}\notin(\Omega\backslash\mathcal{B})$ we get a contradiction from (ii).

If $e^*\in(\Omega\backslash\mathcal{B})$ is such that $(e^*)^T\Lambda e^*>0$ then, since $\kappa\in\mathcal{K}_\infty$, there exists m^* such that for all $m\geqslant m^*$: $-\kappa(m)(e^*)^T\Lambda e^*+\mathcal{W}(x^*,e^*)<0$ which contradicts (B.16).

B.5 Auxiliary lemmas for incremental homogeneity

Preliminarly, recall that $A\preceq B$, $A,B\in\mathbb{R}^{m\times l}$, means $A_{i,j}\leqslant B_{i,j}$ for all $i=1,\dots m$, $j=1,\dots,l$, and $\max_{\theta\in\mathcal{Q}}\Phi(\theta)$, $\Phi\in\mathbf{C}^0(\mathbb{R}^n,\mathbb{R}^{m\times l})$ and compact $\mathcal{Q}\subset\mathbb{R}^n$, represents any matrix M such that $\Phi(\theta)\preceq M$ for all $\theta\in\mathcal{Q}$. If $\{\mathcal{Q}(c)\}$ is a family of compact sets $\mathcal{Q}(c)\subset\mathbb{R}^n$ for each $c>0$ and such that $\mathcal{Q}(c)\to\{0\}$ as $c\to 0$ then $\max_{\theta\in\mathcal{Q}(c)}\Phi(\theta)$ is implicitly assumed to satisfy $\max_{\theta\in\mathcal{Q}(c)}\Phi(\theta)\to\Phi(0)$ as $c\to 0$.

The proof of the following lemma is sketched for reasons of space and follows from simple but lengthy matrix algebra.

Lemma B.4 *Let Φ and Ψ be as in(**H1**). For each $c>0$ and positive definite diagonal $\Gamma\in\mathbb{R}^{n\times n}$ define*

$$N_g:=A+(I+\Gamma A^T)A\sum_{j=1}^{n-1}(\Gamma A^T)^j \tag{B.17}$$

$$M_g(c):=2(I+\Gamma A^T)\Phi_g(c)\sum_{j=0}^{n-1}(\Gamma A^T)^j \tag{B.18}$$

$$\Phi_g(c):=\max_{x',x''\in[-c\mathbf{1},c\mathbf{1}]}\Phi(x',x'') \tag{B.19}$$

$$S_g(c):=2C^TC\Gamma C^T\Psi_g(c)\sum_{j=0}^{n-1}(\Gamma A^T)^j \tag{B.20}$$

$$\Psi_g(c)=\max_{x',x''\in[-c\mathbf{1},c\mathbf{1}]}\Psi(x',x'') \tag{B.21}$$

There exist $\bar{c}>0$ and Γ such that for all $c\in(0,\bar{c}]$

$$-2\Gamma+N_g+M_g(c)+S_g(c)+(N_g+M_g(c)+S_g(c))^T\leqslant-I.$$

PROOF. (*Sketch*). By assumption(**H1**) $\Phi(0,0)=BF$ and $\Psi(0,0)=\alpha C$, $\alpha\in[0,1)$, so that $M_g(0)=2(I+\Gamma A^T)BF\sum_{j=0}^{n-1}(\Gamma A^T)^j$ and $S_g(0)=2\alpha C^TC\Gamma C^TC$ (since $CA^T=0$: see section B.1). Using the fact that $\alpha\in[0,1)$, find positive definite diagonal $\Gamma\in\mathbb{R}^{n\times n}$ such that

$$-2\Gamma+N_g+M_g(0)+S_g(0)+(N_g+M_g(0)+S_g(0))^T\leqslant-2I.$$

The above matrix inequality can be satisfied recursevely on the principal minors with increasing dimensions using Sylvester 's criteria for negative definite matrices and by selecting first $[\Gamma]_{n,n}>0$ up to $[\Gamma]_{1,1}>0$. Finally, pick $\bar{c}\in\mathbb{R}_>$ such that for all $c\leqslant\bar{c}$

$$M_g(c)-M_g(0)+S_g(c)-S_g(0)+(M_g(c)-M_g(0)+S_g(c)-S_g(0))^T\leqslant I$$

taking into account that $\Phi_g(c)\to\Phi_g(0)$ and $\Psi_g(c)\to\Psi_g(0)$ as $c\to 0$.

The following lemma can be proved by using extensively the definition and properties of incremental homogeneity with the properties of the saturation functions (the proof is omitted for lack of space).

Lemma B.5 *Assume* (**H1**) *and let* G *and* Γ *be as in (11).*

(i) A^TAGA^TA *is i.h. with quadruple* $(\mathfrak{r}, \mathfrak{r}+\mathfrak{g}, \mathfrak{g}, A^TA\Gamma A^TA)$,

(ii) $(I-GA^T)\phi$ *(resp.* ϕ*) is i.h.u.b. with quadruple* $(\mathfrak{r}, \mathfrak{r}+\mathfrak{g}, \mathfrak{g}, (I+\Gamma A^T)\Phi(x', x''))$ *(resp.* $(\mathfrak{r}, \mathfrak{r}+\mathfrak{g}, \mathfrak{g}, \Phi(x', x''))$*),*

(iii) $A + (I - GA^T)A\sum_{j=1}^{n-1}(GA^T)^j$ *is i.h.u.b. with quadruple* $(\mathfrak{r}, \mathfrak{r}+\mathfrak{g}, \mathfrak{g}, A + (I+\Gamma A^T)A\sum_{j=1}^{n-1}(\Gamma A^T)^j)$,

(iv) $(I-GA^T)^{-1}$ *(resp.* $I-GA^T$*) is i.h.u.b. with quadruple* $(\mathfrak{r}, \mathfrak{r}-\mathfrak{g}, \mathfrak{g}, (I-\Gamma A^T)^{-1})$ *(resp.* $(\mathfrak{r}, \mathfrak{r}-\mathfrak{g}, \mathfrak{g}, I+\Gamma A^T)$*),*

(v) for each $h \in \mathbb{R}^n_>$, σ_h *(resp.* $\phi \circ \sigma_h$*) is i.h.u.b. with quadruple* $(\mathfrak{r}, \mathfrak{r}-\mathfrak{g}, \mathfrak{g}, 2I)$ *(resp.* $(\mathfrak{r}, \mathfrak{r}+\mathfrak{g}, \mathfrak{g}, \Phi(x', x''))$*).*

References

[1] A. B. Alexandru, S. Pequito, A. Jadbabaie and G. J. Pappas, Decentralized observability with limited communication between sensors, *55th Conf. on Dec. and Contr.*, Las Vegas, NV, 2016, pp. 885-890.

[2] S. Battilotti, Incremental generalized homogeneity, observer design and semiglobal stabilization, *Asian Journ. Contr.*, **16**, 2014, pp. 498-508.

[3] Han W., Trentelman H. L., Wang Z., Shen Y., A simple approach to distributed observer design for linear systems, *IEEE Trans. Autom. Contr.*, DOI 10.1109/TAC.2018.2828103.

[4] M. Deghat, V. Ugrinovskii, I. Shames, C. Langbort, Detection of biasing attacks on distributed estimation networks, *55th Conf. on Dec. and Contr.*, Las Vegas, NV, 2016, pp. 2134-2139.

[5] Li Z., Duan Z., *Cooperative control of multi-agent systems: a consensus region approach*, CRC Press, 2014.

[6] K. Liu, H. Zhu and J. L, Cooperative Stabilization of a Class of LTI Plants With Distributed Observers, *IEEE Tran. Circuits and Systems I: Reg. Papers*, vol. 64, no. 7, pp. 1891-1902, 2017.

[7] L. A. Low, G. Reinhall, D. W. Storti, E. B. Goldman, Coupled van der Pol oscillators as a simplified model for generation of neural patterns for jellyfish locomotion, *Struct. Control Health Monit.*, **13**, 2006, pp. 417-429.

[8] T. Kim, H. Shim and D. D. Cho, Distributed Luenberger observer design, *55th Conf. on Dec. and Contr.*, Las Vegas, NV, 2016, pp. 6928-6933.

[9] M. Mekhail and S. Battilotti, Distributed estimation for feedback-linearizable nonlinear systems, *Eur. Contr. Conf.*, Aalborg, 2016, pp. 2521-2526.

[10] A. Mitra and S. Sundaram, An approach for distributed state estimation of LTI systems, *54th Ann. Allerton Conf. on Comm., Contr. and Computing*, Monticello, IL, 2016, pp. 1088-1093.

[11] A. Mitra and S. Sundaram, Byzantine-Resilient Distributed Observers for LTI systems, *Automatica*, arXiv: 1802.09651v2.

[12] P. Millan, L. Orihuela, C. Vivas, F. R. Rubio, Distributed consensus-based estimation considering network induced delays and dropouts, *Automatica*, **48**, 2012, pp. 2726-2729.

[13] Park S., Martins N. C., , Design of distributed LTI observers for state omniscience, *IEEE Trans. Autom. Contr.*, **62**, 2017, pp. 561-576.

[14] R. Olfati-Saber , R. M. Murray Consensus problems in networks of agents with switching topology and time delays, *IEEE Trans. Autom. Contr.*, **49**, 2004, pp. 1520-1533.

[15] R. Olfati-Saber and R. M. Murray, Consensus problems in networks of agents with switching topology and time-delays, *IEEE Trans. Autom. Contr.*, vol. 49, no. 9, pp. 1520-1533, Sept. 2004.

[16] S. Park and N. C. Martins, Necessary and sufficient conditions for the stabilizability of a class of LTI distributed observers, *51st Conf. on Dec. and Contr.*, Maui, HI, 2012, pp. 7431-7436.

[17] Ren W., Beard R. W., Consensus seeking in multiagent systems under dynamically changing interaction topologies, *IEEE Trans. Autom. Contr.*, **50**, 2005, pp. 655-661.

[18] V. Ugrinovskii, Distributed robust estimation over randomly switching networks using $\mathcal{H}_\infty$ consensus, *Automatica*, arXiv: 1512.01294v1.

[19] H. Zhu, K. Liu, J. L, Z. Lin and Y. Chen, On the cooperative observability of a continuous-time linear system on an undirected network, *Int. Joint Conf. on Neural Networks*, Beijing, 2014, pp. 2940-2944.

[20] Yu W., Chen G., Cao M., Kurths J., Second-order consensus for multiagent systemswith directed topologies and nonlinear dynamcis, *IEEE Trans. Syst.Man Cybernetics, Part B*, **40**, no. 3, 2010, pp. 881-891.

[21] L. Wang and A. S. Morse, A Distributed Observer for a Time-Invariant Linear System, *IEEE Trans. Autom. Contr.*, 2017 (10.1109/TAC.2017.2768668).

[22] L. Wang, A.S. Morse, D. Fullmer, J. Liu, A hybrid observer for a distributed linear systems with a changing neighbor graph, arXiv: 1706.04235v1.